Journal of
Experimental Psychology

VOL. VII., No. 3 June, 1924

STUDY OF ANIMAL DRIVES

BY FRED A. MOSS

George Washington University

INTRODUCTION

Basic Thesis: *The behavior of any animal is the resultant of his drives to action and the opposing resistances.*

It is a fact of common observation that animals when confronted by certain situations tend to respond with rather definite, fixed, behavior. It is to these tendencies to action that I apply the terms 'drive' and 'resistance.'

Drive.—By 'drive' I mean the impelling forces in the situation that stimulate the animal toward certain positive behavior. For example, if an animal is kept without food for a certain length of time, certain organic stimulations of the nerve-endings in the stomach provoke in the animal restless and seeking behavior until food is found and the "drive" for the time being, stopped. It is to the motive-force behind this seeking behavior that I apply the term 'drive.'

Resistance.—By 'resistance' I mean the repelling forces in the situation which stimulate the animal to negative reactions, or tend to draw him away from certain behavior. For example, the mere presence of cats produces in rats certain 'resistances,' and these 'resistances' stimulate the rats to behavior tending to keep, or remove, them from a closer contact with cats.

EXPERIMENTAL PROCEDURES AND RESULTS

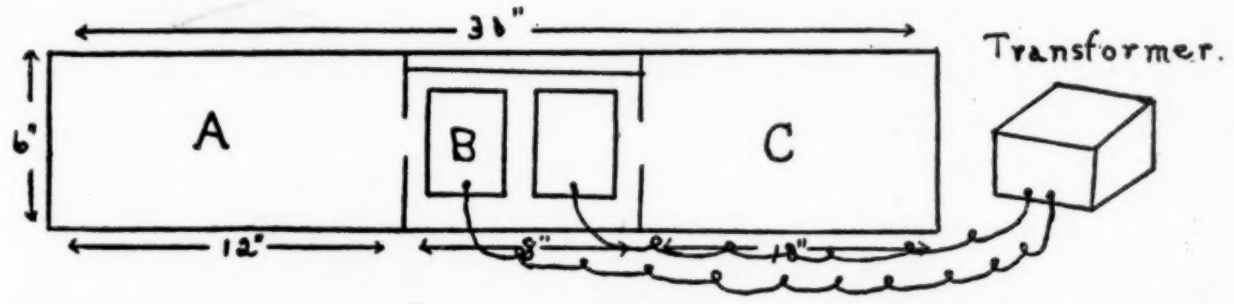
The animals used were white rats, ranging in ages from one hundred twenty-five to two hundred days. All males are distinguished by odd numbers and all females by even numbers.

Preliminary Training.—In a series of preliminary experiments in which the electric shock was used to produce the opposing stimulus, it was found that on the first trial the rat walked with little hesitation on to the brass plates, and when he had received the first shock he continued across the plates about as often as he turned back. However, in his subsequent experiences with the plates his behavior was quite different. As a rule, unless given some very strong stimulus which would make him do otherwise, he tended to keep clear of all contact with the plates. When given the necessary stimulus to make him approach the plates, he would go up to them very slowly and would cautiously reach out his foot to touch them. In most instances the foot was jerked back before it had even touched the plates. After crossing the plates and receiving the shock the second time, this avoiding behavior was more pronounced. I see no better way of explaining this change in the animal's behavior than to say that from his first two experiences with the plates he had "learned" to avoid them. From these experiments and from many other observations, I am convinced that with the lower animals as well as with the human, a single experience is very often sufficient to cause a marked change in the animal's future behavior. Take the following case as a typical illustration of my meaning: A rat stepped on the plates and received an electric shock of ninety volts. The effect of that single experience was sufficiently strong that when the animal was kept without food for three days and placed in the experiment box, with food on the opposite side of the plates, he would not touch the plates nor try to cross them to get to the food. Instead he tried to get to the food by climbing the sides of the box, and failing in this, he tried other methods equally futile. Altho left in the box for two hours, during which time he was almost constantly active, he always carefully avoided the only route that could bring success, the straight, direct route over the plates. Just as in the case of the man who having gone thru certain experiences is 'never the same again,' so this rat, having once received the shock, would never willingly come in contact with the plates again.

In order to make doubly sure that from the first of each experiment the brass plates might serve as a repelling stimulus, all animals used in the experiments were given two experiences with the plates before the actual experiment began. In these preliminary experiences they received a shock of 28 v. (twenty-eight volts).

I. The first experiment was devised to measure in terms of a constant resistance the strength of hunger drives varying in duration from twelve to one hundred forty-four hours.

In this experiment and in most of the others reported, I used an experiment box, the floor-plan of which is shown in the following diagram.



It was a simple box, 30 inches long, 6 inches wide and 12 inches deep. All sides except the front were made of wooden boards. The front side was made of wire fly-screen netting. It was divided into three sections, *A*, *B*, and *C*, the partitions being at *x* and *y*. The partitions were made of ¼ inch rat wire and equipped with doors of the same material. Section *A* was the full width of the box and twelve inches long. It was so made that it would hold water to the depth of one inch. When using the electric shock as an opposing stimulus, a small quantity of water was always kept in this section. On their way to the brass plates the rats passed thru this water and thus insured an almost perfect and uniform contact. Section *B* was eight inches long, but by nailing a board on the back side, the width was made five instead of six inches. In this section was placed the brass plates, so arranged that when the rat touched them he received an electric shock. In order to prevent the rat's trying to jump the plates or climbing the wire along the front, a wire screen was stretched about three inches above the plates, from partition *x* to partition *y*. This made it impossible for a rat to go from *A* to *C* without touching

the plates and getting an electric shock. The amount of current was controlled by a step-up transformer. Section *C* was ten inches long and the full width of the box, In this section the food was placed.

In this experiment sixteen rats were used, eight of each sex. They were divided into four groups of four each. Those in the first group were kept without food for twelve hours, those in the second group for twenty-four hours, those in the third group for forty-eight hours, and those in the fourth group for seventy-two hours before being given a trial in the experiment box. An electric shock of 20 v. was used for the stimulus opposing the drive. Each rat was placed in section *A* and left for one hour unless he crossed to the food in section *C*.

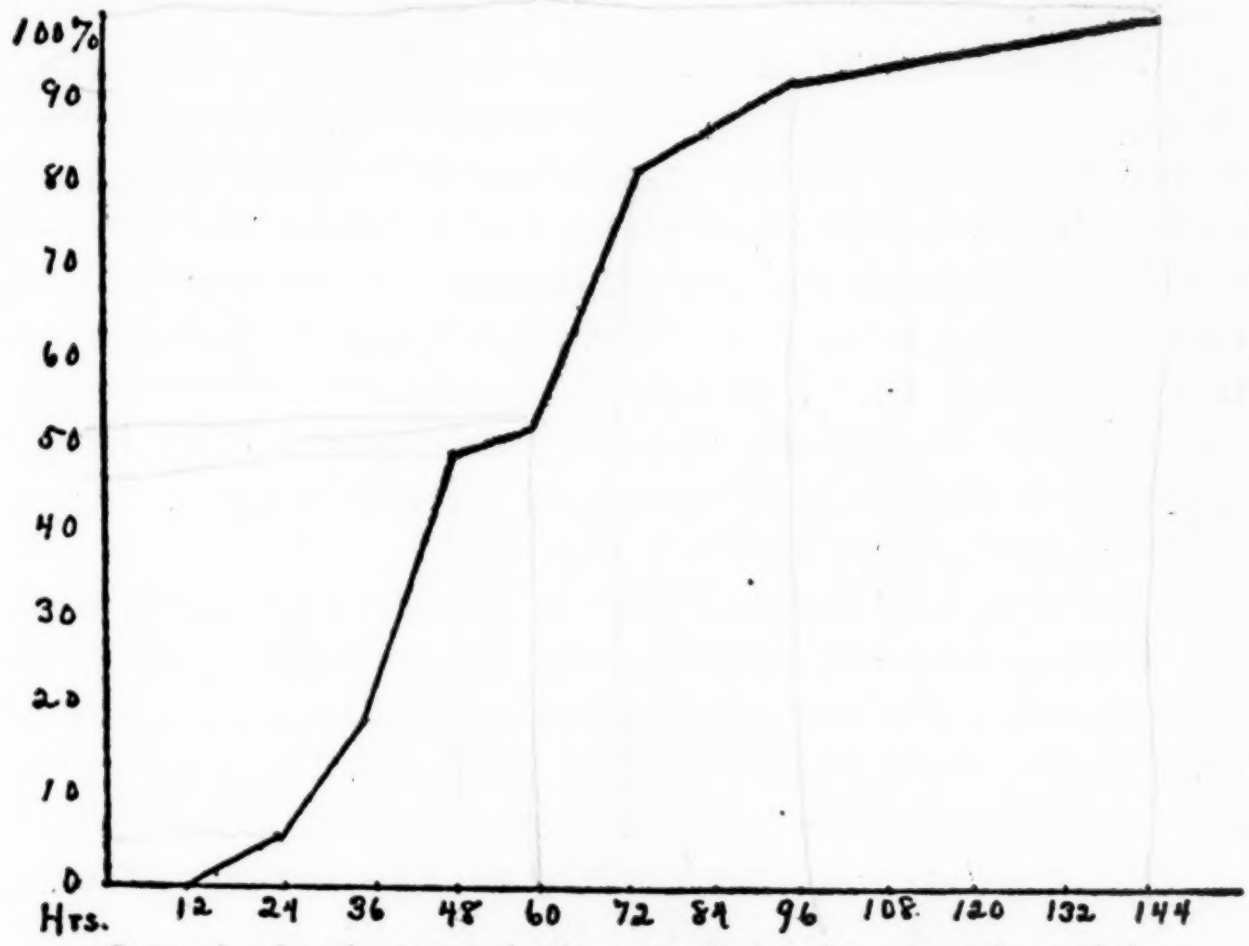
None of the animals in the first group, No. 1, 2, 3, and 4, crossed in the time allowed. While in the second group, composed of rats Nos. 5-8, only one, No. 6, crossed. In the third group, rats 9-12, three crossed, Nos. 10, 11, and 12. In the fourth group, 13-16, all crossed.

The eight rats that had failed to cross, Nos. 1, 2, 3, 4, 5, 7, 8, and 9, were kept without food for another twenty-four hours and placed in the box again. On this trial four crossed, Nos. 1, 2, 7, and 8, The remaining four were again deprived of food for another twenty-four hours and the procedure repeated.

On this trial two crossed, Nos, 3 and 5. After the remaining two had been starved for an additional twenty-four hours, they were given another trial in the box and both crossed.

The activity of each rat on each trial was scored from 1 to 6, almost constant inactivity being scored 1, and practically continuous activity was scored 6. The average activity score with a twelve-hour hunger drive was only 2. Of those with a twenty-four-hour drive, the score was 3; with the thirty-six-hour drive the score was 3.5; with the forty-eight-hour drive the score was 3.75; with the sixty-hour drive the score was 4.5. With both the seventy-two and eighty-four-hour drives the score was 5. While with the ninety-six and the one hundred forty-four-hour drives the score dropped to only 3. This last low score was due to the inactivity of a large male, that was very sullen and inactive.

The following graphs show the percentages having crossed at the various starvation periods, also their activity score at these periods.



Curve showing the percent having crossed at various starvation periods.

Curve showing activity at various starvation periods.

II. After having determined the effects of varying hours of hunger, it appeared desirable to measure the comparative strength of a 72 hour hunger drive and a 28 v. resistance.

In the first experiment the resistance, or opposing stimulus, was kept constant at 20 v. and the strength of the drives was raised until *all* the animals overcame the resistance. While in the experiment that we are now considering, the strength of the drives was kept constant, being determined by a 72 hour starvation period, and the strength of the resistance by an electric current of 28 v. The same apparatus was used as in experiment one. Ten animals were used, five males and five females, Nos. 17-26. Each animal was placed in section *A* and left for one hour unless he crossed in the meantime to the food in section *C*.

With a 72 hour hunger drive the animals were very active, the average activity score for this group being 5. All the animals crossed in the time allowed except two, Nos. 22 and 25. These results would indicate that in four cases out of five a 72 hour hunger drive is stronger than a 28 v. resistance.

III. An effort was next made to measure the comparative strength of the sex drive and a 28 v. resistance.

There are many factors that determine the strength of the sex drive, such as age of the animals, their family inheritance, types of food they have been fed, state of health, etc. In a series of preliminary experiments it was discovered that there were very wide individual differences in the strength of the sex drive; one individual would sometimes overcome three times as much resistance to satisfy the sex drive as another. It was also found that there was not nearly so much difference among individuals of the same family as among those of different families. This was true not only in the strength of their drives but in the activity score as well. It was further shown that in the life of the same animal, the sex drive differs materially at different ages; ranging all the way from approximately zero strength during the period from birth up to about sixty-five days of age; then it has a very rapid increase in strength, reaching its highest point about the age of one hundred fifty or one hundred seventy-five days.

In as much as I wished to compare the results of this experiment with the results obtained in the second experiment, and thus gain some idea of the comparative strength of the sex and hunger drives, only those animals were selected for this experiment that had a brother or a sister among the group used in experiment two. Care was taken to get the two groups not only from the same families, but where possible, from the same litter.

The two groups were fed the same kind of food and kept under as nearly the same conditions as possible prior to these experiments.

As in experiment two, ten rats were used, five males and five females, Nos. 27-36. The same apparatus was used as in the first experiment, the only change being that the door to section *C* was closed in order to keep the rats that were used as a stimulus from leaving *C*. Otherwise those in section *C* would have been about as likely to cross to *A* as those in *A* to cross to *C*. It was so arranged that by attaching a string to the door and running it thru a series of pulleys, the door could be opened when an animal from *A* crossed the plates in *B*.

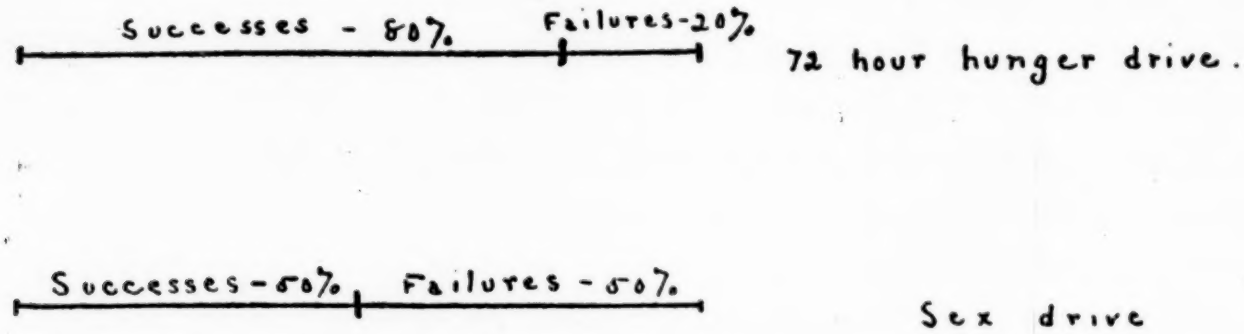
One of the females when in heat was placed in section *C*. The males were placed in section *A*. Only two of the males, Nos. 27 and 33, crossed to the female in the hour's time that they were left in the box.

The complement of this experiment was performed by confining the males in section *C* and leaving the female in section *A* free to cross the plates if she had sufficient drive to overcome the resistance. All the females thus tested were in heat. Three of them. Nos, 30, 32, and 36, crossed the plates in the time allowed.

It is of interest that when impelled by the sex drive only five of the ten overcame a resistance of 28 v., but when impelled by a 72 hour hunger drive eight of the ten overcame the resistance. The following lines will indicate the comparative success and failure of the two drives against the same opposing stimulus of 28 v.

Unless some important factors have been overlooked, these results would tend to disprove the commonly accepted theory that the sex drive is the strongest of all drives.

It is significant that when driven by the sex urge, three of the females crossed and only two of the males. These results might also throw the shadow of suspicion on the cor-



Successes and Failures

rectness of the statement that 'the male is essentially the pursuer and the female the pursued.' Whether the above statement is correct or not, these results certainly show that some females will overcome more resistance to place themselves in a position to be 'pursued' than will the average male in order 'to pursue.' The behavior of the female when admitted to section *C* would, at first glance, tend to support the statement that she is the pursued rather than the pursuer. On the approach of the male she responds with the proverbial 'fleeing' behavior, only she doesn't flee very far nor into inaccessible places such as back across the plates into section *A*. A quotation from Thorndike would very accurately describe this situation:

"The female responds by coy advances and retreats: the male by caressing pursuit and capture. The former is satisfied by, and so instinctively maintains, whatever augments the aggressiveness of the male; he responds similarly to the hopeful difficulties which her behavior offers. Capture and submission are responded to by the mutual absence of fear, . . . and by satisfaction in bodily contact, including as the final element the contact necessary for the fertilization of the ovum."

IV. Having determined the strength of the normal sex drive an attempt was next made to measure the effect of perversions on the strength of this drive.

The same apparatus was used in this experiment as in the preceding one. Care was taken to get animals of the same

age, and in so far as possible, of about the same physical condition as in the two preceding experiments.

Five males, Nos. 37, 39, 41, 43, and 45, were used. These rats had been kept in a cage by themselves from the age of twenty-five days till the time of this experiment, approximately one hundred fifty days. During this time they never came in contact with a female. On several occasions certain perversions of the homosexual type had been observed among them.

They were placed in section *A* in two groups, the first group containing two rats, and the second group the other three. A female in heat was placed in section *C*. The opposing stimulus was 28 v.

Altho left in the box for one hour, not a single one of the rats crossed to the female, nor did they seem to be excited by her presence. They were very inactive, their activity score being only 2. These results are very interesting when compared with those in experiment three. They would indicate that when the animal has found some way, other than the normal one, of satisfying the sex desire, the 'pulling force' of the normal stimuli on that individual is considerably weakened.

V. An effort was next made to measure the strength of the maternal drive.

In this experiment it was assumed that the strength of the maternal drive may be judged by the amount of resistance the mother will overcome in her efforts to go to her young, especially if she hears the distress cries of the young.

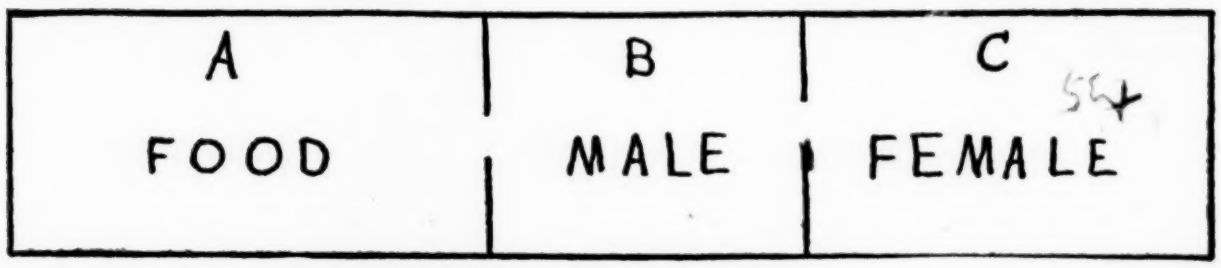
Five females, Nos. 38, 40, 42, 44, and 46, together with their newly born litters, were used in this experiment. The same apparatus was used as in the other experiments.

The mother was placed in section *A* and the little ones in section *C*. The little ones were punched and pushed about to make them squeal. In this way the attention of the mother was called to their distress.

Only one of the five mothers took the shock to get to her young. It happens that she was the wildest one in the group, and had been handled the least. The others were very tame.

It is quite possible that these rats had been handled so much that when they saw the hand of the experimenter punching the young ones, the distress cries lost their effect.

VI. In one of the earlier experiments an effort was made to determine the comparative strength of the sex and food drives by measuring the amount of resistance each would overcome. An attack was also made on the problem by balancing the two drives against each other, as shown in the following:



Food and sex balance.

In this experiment I used the same apparatus that I used in experiment three with the exception that the electrical apparatus was removed from section *B* and the door to section *A* was closed. The doors from section *B* to the other two sections were fixed so that they could be easily opened from section *B* but could not be opened from either of the other two sections, so that the animals could easily get into one of the other sections, but, once in, could not get back into *B*. All the rats used in this experiment had, prior to the experiment, been placed in section *B* and given the opportunity to learn that if they went into either *A* or *C* they were kept one hour before being taken out. All of the five had this experience four times before this experiment began. I gave them this preliminary training because I thought that it might have the effect of making the rats hesitate before going into the section from which he happened to get the first stimulus. This would give him the opportunity of getting the two sets of opposing stimuli from both sections before going into either.

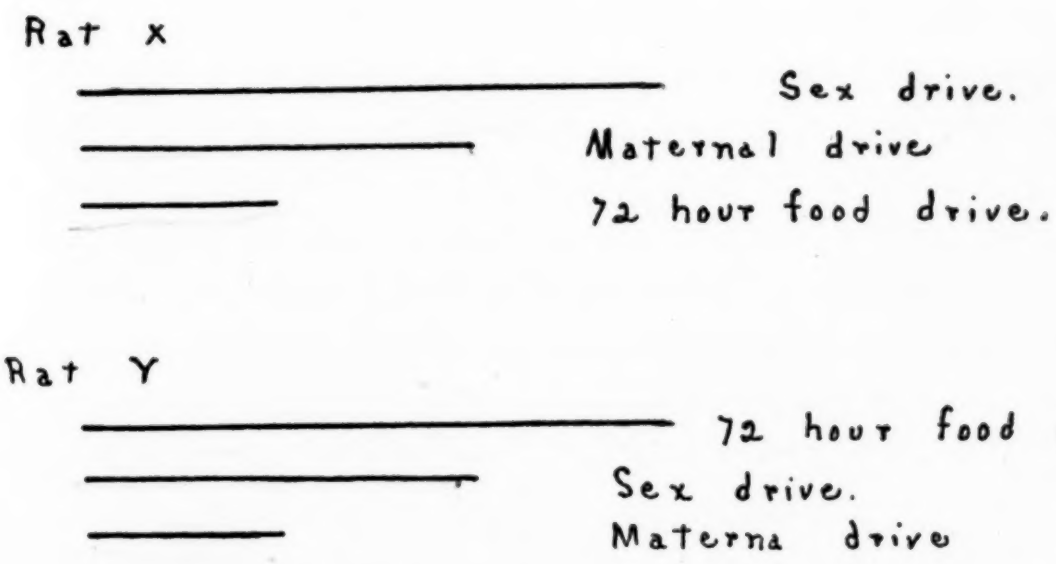
In this experiment the same five males were used as in experiment three, Nos. 27, 29, 31, 33, and 35.

A female when in heat was placed in section *C* and food was placed in section *A*. The males after having been de-

prived of food for 72 hours were placed in section *B* and given the 'choice' between food and sex gratification.

Only one rat, number 27, went into section *C*; the other four went immediately to the food in section *A*. It seemed to take No. 27 a long time to 'make up his mind.' He ran first to one partition then to the other. He would stick his nose thru the wire of one partition then go over to the other and repeat the behavior. If animals have "mental conflicts" I would judge from his behavior that he experienced one. Be that as it may, after being in section *B* thirty-two minutes he went into section *C*.

It is worth noting that number 27 was the rat that in experiment three overcame the resistance of 28 v. to satisfy his sex drive. Of course one case shows little, but from my other observations on rats, I have a feeling that if we were to test a sufficient number of cases, we would very likely find the sex drive in some rats three or four times as strong as it is in others of the same age and sex. At the same time, the rats with the strong sex drive may be weak in other drives. The following lines will illustrate some individual differences in drives. The length of the line indicates the strength of the drive.



Taken uncritically, these results would indicate that the 72-hour hunger drive is considerably stronger than the sex drive. But we should consider that in this experiment we may not be measuring the normal strength of the sex drive. It is quite likely that as the hunger increases the sex drive

decreases. To illustrate, it is conceivable that with a full stomach, the sex drive is equal to 100 units, but with an empty stomach, due to having been deprived of food for seventy-two hours, the strength of the sex drive is only 65 units. A series of experiments to be conducted in the near future has been planned which will tend to clear up this point.

VII. In the next experiment a study was made of the rat's choice between the 'lesser of two evils.' This was done by measuring the amount of decrease in temperature of the water in which the animal was standing, necessary to force the animal to climb on a brass plate where he received a 12 v. shock.

The same apparatus was used as in experiment one with the addition that section *A* was filled with water to the depth of one half inch. Crushed ice and a thermometer were also provided. The transformer was set to turn a 12 v. opposing stimulus on the plates. Section *C* contained neither food nor female.

Ten animals were used. When the temperature of the water was around 25 degrees centigrade, the animal was placed in section *A*. Crushed ice was quickly poured into the water until the temperature of the water was reduced sufficiently to make the animal climb out of it on to the plates. In this experiment we have one resistance overcoming another. It is what the moralists might call 'the choice between the lesser of two evils'; while the man on the street would describe the situation as being 'between the Devil and the deep blue sea.' Regardless of what name is given to the situation, the actual results are shown in the following tale.

Temperature	Cases Climbing Out
20–25° C	0.
15–19 C	1.
10–14 C	1
5– 9 C	5.
1– 4 C	3

VIII. The next experiment was planned to determine the amount of reinforcement necessary to make a twelve-hour hunger drive overcome a 28 v. resistance.

Here we have a situation in which the resistance is very much stronger than the drive. In the first experiment we found that a 12-hour hunger drive was in no instance strong enough to overcome a 28 v. resistance. In the present experiment we are trying to add to this drive an antagonistic resistance that will make it able to overcome the opposing resistance.

The same apparatus was used as in the last experiment with the addition of food in section *C*. The transformer was set for 28 v.

Ten animals, five males and five females, were used. The animals having been deprived of food for twelve hours were placed in section *A* and crushed ice was fed into the water until they crossed to section *C*. The temperature of the water was recorded at the time each one crossed.

The following table shows the amount of reënforcement necessary to make the twelve-hour hunger drive overcome the 29 v. resistance.

Temperature	Number Crossing
20–25° C	0 ✓
15–19 C	2 ✓
10–14 C	4 ✓
5– 9 C	2
0– 4 C	2

IX. In the next experiment an effort was made to measure the effect of having overcome a resistance upon the strength of the resistance.

In this experiment I used the same apparatus and the same animals that were used in the last experiment. For five successive days the animals were tested under approximately the same conditions as in the last experiment. All animals were kept without food for twelve hours before testing. The resistance, or opposing stimulus, was kept at 28 v.

On the first trial the temperature of the water had to be lowered on the average to ten degrees *C*. before the animals crossed. On the second trial the average temperature at which the animals crossed was two degrees higher, 12° C. On the third trial it went up to 13.5° C. While on the fourth there was no improvement, in fact a slight falling off over the

last trial. On this trial the average temperature was approximately 13° C. On the fifth trial there was an improvement, the average temperature being raised to 14° C.

Thus we see that five trials reduced the amount of reenforcement necessary to make a twelve-hour drive overbalance the 28 v. resistance, by about four degrees C. From these results it is reasonable to conclude that a resistance tends to lose its repelling force each time it is overcome.

X. It is a fact of common observation that when an act is performed and followed immediately by punishment that that act is less likely to be performed again. A series of experiments were devised to measure the effect of punishment on the rats' hunger drive for cheese.

In one corner of a simple cage was placed a block of cheese, three inches long and one inch in width and thickness. The approach to this was so guarded by brass plates that every time the rat tried to nibble the cheese, he got an electric shock of 60 v.

In this experiment five rats were used. The rats were kept supplied with a liberal amount of food, other than cheese, so that they could not become very hungry. The animals were left in the cage with the cheese thus protected for one week. When the protecting electrical apparatus was removed, none of the animals disturbed the cheese, although it was left in the cage unprotected for a week. From these results we are led to conclude that by punishment we can build up in an animal a permanent inhibition of a desire for a particular kind of food.

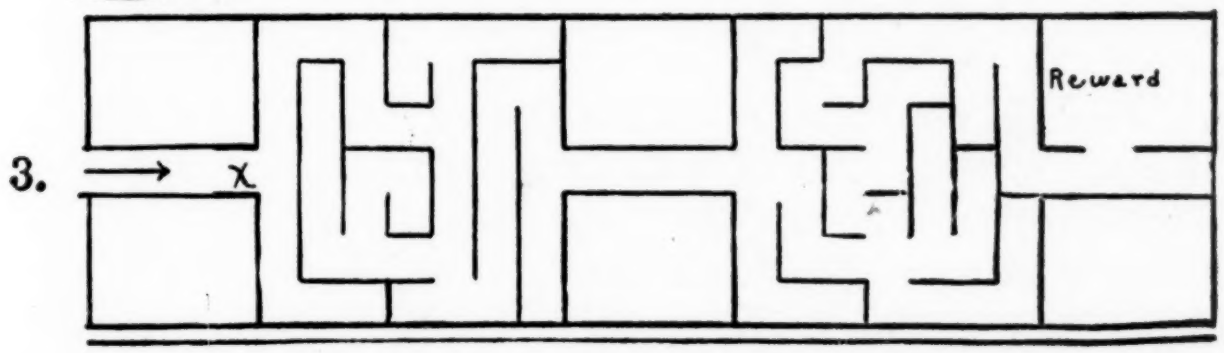
Prof. J. J. B. Morgan, in an unpublished experiment, has studied the effect of punishment on the sex drive. He placed three males and three females in adjoining cages separated only by a wire partition and left them for several weeks. Every time they tried to get in contact with each other they got an electric shock. Among his other findings is the interesting case of a female whose normal sex behavior was so disturbed that ever afterwards, she fled in terror at the mere approach of the male.

XI. From time immemorial a battle has been raging in the field of education concerning the comparative worth of punish-

ment and reward as incentives in learning. With the hope that some light might be thrown on this problem from the field of animal psychology, an experiment was devised to measure the comparative efficiency of the hunger drive, the sex drive, and the resistance to cold, as motive forces in learning a maze.

There are three elements that must be considered in attempting to measure the efficiency of learning methods: the time it takes to learn the problem; the number of errors made in learning it; and the durability, or permanancy of the learning.

The maze used in this experiment was of the same general pattern as number three of the Army Beta Intelligence test. All the alleys were four inches wide and all other parts were extended in proportion. The following diagram needs little explanation.



The rats were placed in the maze at "*X*." In the compartment marked "Reward" was placed the food or the female while trying the first two sections of this experiment; when working on the third section of the experiment, this part was the only part of the maze that did not contain ice.

Fifteen animals were used, five for each part of the experiment. They were about the same age, and special care was taken to get the groups as similar as possible.

The five animals in the first group were kept without food for 48 hours and placed in the maze one at a time. The time and errors were recorded for each. The average time taken by the five in this group to find the food on the first trial was 38 minutes with 29 errors. They were permitted to eat for fifteen minutes, and then removed from the maze. They were given another trial the next day. This time they made

it in 21 minutes with only 17 errors. They were again permitted to eat for fifteen minutes and removed.

This procedure was repeated for the next three days. The average time for the five on the 3d, 4th, and 5th trials was in 3 min. 50 sec. and the average number of errors was 7.

In the second part of the experiment where the sex stimulus was used, five males were placed in the cage one at a time and a female in heat was placed in the 'reward' box. The time and errors were recorded as in the first part. When the male reached the reward box he was left in it with the female for fifteen minutes. The average time of the first trial for the rats to find the reward box was 49 minutes with 31 errors. On the second trial the average time was 22 min. 14 sec., with 19 errors. The average time for the 3d, 4th, and 5th trials was 3 min. and 5 sec., with 8 errors.

In the third part of the experiment where punishment was used as the incentive, the floor of all parts of the maze except that of the 'reward' box was covered with crushed ice. The only way the animals could escape from the cold was to go to the reward box. The average time on the first trial was 36 minutes with 29 errors. On the second trial the average time was 28 minutes with 18 errors. And the average for the 3d, 4th, and 5th trials was 5 minutes and 20 sec. with 12 errors.

The results of the three parts of the experiment are combined in the following table.

TIME TABLE

Trial	Food	Sex	Punishment
1st	38 min.	49 min.	36 min.
2d	21 "	22m. 14 s.	28 "
3d, 4th, 5th	3 m. 50 s.	3 m. 5 s.	5 m. 20 s.

ERRORS TABLE

Trial	Food	Sex	Punishment
1st	29	31	29
2d	17	19	18
3d, 4th, 5th	7	8	12

Thirty days later these tests were repeated to measure the permanency of the learning. Each rat was given two trials.

The average results of these two trials are shown in the following table:

	Food	Sex	Punishment
Time	6 min. 40 sec.	5 min. 12 sec.	10 min. 18 sec.
Errors	16	13	17

These results would indicate that for immediate efficiency it makes little difference whether you use the sex or food rewards. Both seem superior as incentives to this particular punishment method. When the permanency factor is considered it seems that the sex reward works some better than that of food; and that the punishment method is much inferior to both of these.

General Discussion of Experiments

In the foregoing experiments I have endeavored to set forth a method whereby some of the so-called 'intangible' elements of animal behavior can be objectively measured. This method enables us to evaluate the 'motives,' or impelling forces, in behavior. In the past, students of animal behavior have been too content to study the mere act and have paid little attention to the impelling forces in the act. While the methods used here are admittedly very crude and imperfect, it is hoped that they will serve as a starting point for further investigations which will devise a greatly improved technique to replace them.

There was a time in the history of the race when our forefathers were very pleased to be able to measure things in pounds, for it was so much more accurate than guessing. They were even more pleased and certain of their results when they got down to ounces. The ounce as a fine unit of measure has yielded place to the grain, and the grain in turn has been replaced by still finer units of measure and they by still finer and more accurate, until to-day the U. S. Bureau of Standards has perfected an instrument that will measure accurately down to the 1/50,000,000 (one fifty millionth part of an ounce). The instruments used in these experiments may not even be perfected to the pound stage of accuracy, but they at least get

away from the "popular guessing contests" which have been characteristic of the various attempts in the last twenty-five years to find 'the one' impelling force, or instinctive drive behind all animal behavior. A current example of such a 'guessing contest' is the attempt of some of the over-enthusiastic disciples of Freud to explain everything in terms of the sex drive. By the methods described in this paper we can prove that the animal has not only many other drives and resistances which play a very large part in determining his behavior, but we can show that some of these other drives are even stronger than the sex drive. Take for example the results obtained in experiments three and six. These two experiments objectively prove that a situation caused by keeping the animal without food for seventy-two hours generates in him a stronger drive than will the situation caused by the presence of a female in heat.

It is no longer a matter of opinion as to the strength of a steel beam: Construction engineers can very accurately measure the amount of stress and strain that it can withstand. They have been able to do this by substituting numerical facts for opinions, reliable units of measure for mere guess work. The methods used in these experiments at least yield their results in numerical measures, however crude and coarse they may be. With the perfection of these methods it is reasonable to expect that the human engineer will also be able to measure the strength of the materials with which he deals, and to tell with a high degree of accuracy the amount of stress and strain an individual may be expected to withstand.

Conclusions

The primary object in performing these experiments was to develop a method for numerically measuring drives and resistances. The experimenter was only incidentally interested in reaching conclusions as to the relative strength of the various drives and resistances. The only way to be certain that the average performance of a group is typical, or thoroughly representative, of what you might expect from that species, is to have a fairly large number of chance-selected

animals in the group. While great care was taken to eliminate all apparent selection, yet, owing to the small number of animals in the various groups, it is possible that in some cases the element of selection entered. This should be kept in mind when making any sweeping conclusions from this data. This precaution should be remembered while reading the following conclusions which these experiments seems to justify.

1. Experiment one shows that a resistance may be overcome by increasing the strength of the drive. It also shows, as do most of the other experiments, that vast individual differences exist in the strength of the different drives and resistances.

2. Experiment two shows that if a 72-hour hunger drive is opposed by a 28 v. opposing stimulus, or resistance, in eight cases out of ten the hunger drive will overcome the resistance. A 72-hour drive would then seem to be about the threshold for overcoming this resistance.

3. Experiment three shows that a 72-hour hunger drive is in most cases stronger than the sex drive. It also shows that the female will overcome as much, or more, resistance to secure sex gratification, as will the male.

4. Experiment four shows that perversions tend to very greatly weaken the strength of the normal sex drive.

5. Experiment five tends to show that the maternal drive is not as strong as the sex or 72-hour food drive.

6. Experiment six shows that with most animals the 72-hour food drive is stronger than the sex drive. It also shows that the sex drive is in some animals much stronger than it is in others. Take for example the case of rat No. 27.

7. Experiment seven shows that one can make an animal choose something which is very repulsive by making the other possible choice more repulsive. (I believe most suicides show this overcoming of a natural resistance by a still stronger resistance.)

8. Experiment eight shows that although a drive may be much weaker than a resistance, it can be reënforced by other drives or antagonistic resistances until it overcomes the resistance.

9. Experiment nine shows that overcoming a resistance tends to weaken the strength of the resistance.

10. Experiment ten shows that we can greatly modify and possibly break down a natural drive by punishment.

11. Experiment eleven indicates that rewards are better than punishments as motive forces in learning such problems as the maze.

From the results obtained in these experiments, I believe that we are justified in stating the following general theses which may be used in the interpretation of any animal behavior, whether the behavior be that of a rat, dog, horse or human.

1. Any animal drive may be measured in terms of the resistance overcome, provided the strength of the resistance is known. Or, where the strength of the drive is known and that of the resistance is not, the resistance may be measured in terms of the drives.

2. Any drive that succeeds is stronger than the resistance overcome, and any drive that fails is not as strong as the resistance. For example, if a hunger drive caused by a forty-eight-hour starvation period, will make the animal take an electric shock of 12 v. to get to food, but fails to get him to take the shock of 20 v., we may say that the strength of a forty-hour hunger drive lies between a 12 v. and a 20 v. opposing stimulus, being stronger than the 12 v. but weaker than the 20 v.

3. If two independent drives are opposed by the same resistance and one drive overcomes the resistance while the other fails, the one that overcomes the resistance is the stronger drive. For example, if a seventy two-hour-hunger drive overcomes a 20 v. resistance and a sex drive fails to overcome the same resistance we may conclude that the hunger drive is stronger than the sex drive.

4. Given two drives both functioning at the same time, and so arranged that neither can succeed without neglecting the other, the one that succeeds is the stronger drive. This is well illustrated in the balancing of the hunger drive against the sex drive in experiment six.

5. Given two antagonistic resistances both functioning at the same time and both so arranged that neither can cease to function without overcoming the other, the one that overcomes is the stronger resistance. This is illustrated in experiment seven, where the two opposing stimuli are the ice and the electric shock.

6. When one drive by itself is not strong enough to overcome a resistance, it may be reënforced by other drives, until it is strong enough to overcome the resistance. Experiment eight is a very good illustration of this thesis.

7. As a motive force in provoking the learning of such problems as the maze, that drive is strongest which causes the animal to learn the problem in the shortest time and with the fewest errors. This was illustrated in experiment eleven.

8. Other things being equal, every time a resistance is overcome, the strength of that resistance is weakened. This is illustrated in experiment nine.

9. The combined results of these eleven experiments seem to substantiate the basic thesis that *the behavior of any animal is the resultant of his drives to action and the opposing resistances.*

Bibliography

1. Donaldson, H. H. 1915. The Rat. Philadelphia.
2. Hubert, Helen B. 1914. Time versus Distance in Learning. *J. Animal Behavior*, Vol. 5, pp. 66–72.
3. Ishii, O. 1922. Observations on the Sexual Cycle of the White Rat. *The Anatomical Review*, Vol. 23.
4. Long, A. J. and Evans, H. M. 1922. The Œstrous Cycle in the Rat and its Associated Phenomena. Memoirs of University of Cal., Vol. 6.
5. Moss, Samuel. 1836. Notes on the Habits of the Domesticated White Rat and the Terrier Dog that Lived in Harmony Together. London Mag. Nat. History, Vol. 9.
6. Small, W. S. 1900. An Experimental Study of the Mental Processes of the Rat. *Am. J. Psychology*, Vol. II.
7. Thorndike, Edward L. 1898. Animal Intelligence. Macmillan Co, New York.
8. Vincent, S. B. 1915. The White Rat and the Maze Problem. *J. Animal Behavior*, Vol. 5.
9. Watson, John B. 1914. Behavior. An Introduction to Comparative Psychology. Henry Holt & Co., N. Y.
10. Yerkes, R. M. 1907. The Dancing Mouse. Macmillan Co., N. Y.

THE EFFECT OF TRAINING ON INDIVIDUAL DIFFERENCES

BY H. B. REED

Grinnell College, Grinnell, Iowa

The effect of equal practice or equal education on individual differences has been stated very definitely by our leading psychologists. Thorndike, for example, says: "The facts found are rather startling. Equalizing practice *seems to increase differences*. The superior man seems to have got his present superiority by his own nature rather than by superior advantages of the past, since during a period of equal advantages for all, he increases his lead." *Educational Psychology*, Vol. III., page 305.

"All experimental results point in the direction that practice does not equalize abilities; in fact, equal practice tends to increase differences in achievment and skill rather than to decrease them. The more gifted individuals profit more, both relatively and absolutely, than the less gifted. This experimental fact is one of the most profound bits of evidence regarding the whole problem of heredity and environment. The talented men not only start with greater initial capacities but seem also to be capable of more intense application and more zealous desire to improve. 'To him that hath shall be given' is psychologically true in the sphere of intellectual training as well as in the sphere of morality and religion. The man with ten intellectual talents will acquire far more than the man with one talent. If we may generalize for life as a whole, equal opportunities for all do not produce equal abilities in all. Men may be born free politically, but they are not born equally mentally; they may be born equal in opportunities in a democratic society, but they are certainly not equal in their ultimate achievements in life."—Starch, *Educational Psychology*, 1918, page 91.

"While it is, therefore impossible to state adequately or definitely the amounts and rate of gain by either a gross gain

or a per centile method, an inspection of the table shows that by any rational method of interpretation the scriptural law 'to him that hath shall be given' holds. Those with the highest initial scores have the highest final scores and gain most absolutely and relatively."—Henmon, *Journal of Educational Research*, Vol. I., page 94.

After reading such excellent authorities one is tempted to leave the matter rest, and believe that here is a question to which science has found a definite answer, but our peace of mind is disturbed as soon as we begin to examine the logic and the measurements upon which these statements are based. If we measure improvement by increase in the amount of work per unit of time and individual differences by the absolute amounts of the increase, then the above statements are justified, but as soon as we measure improvement by decrease in the amount of time per unit of work and individual differences by the absolute amounts of the decrease, then we are justified in drawing diametrically opposite conclusions. I presume that all will agree that amount of time per unit of work is just as accurate a way of measuring efficiency as amount of work per unit of tine. But the two methodds lead to opposite conclusions about the effect of equal practice on improvement and individual differences. It is the purpose of this paper to show that since this is the case, neither of these methods are reliable, that the above quotations are therefore false, and that when individual differences are measured by more accurate methods we must conclude that equal practice decreases individual differences.

First allow me to illustrate how the two aforementioned methods lead to opposite conclusions. *A* and *B* practice 30 minutes in addition. Improvement is measured by the gain in in the number of examples from the first minute to the last minute. Table I. gives the results.

Conclusions.—*B*, who had the highest initial performance, gained much much more than *A*. An equal amount of practice made them more different. Hence although individuals may be born with equal opportunities, heredity makes them widely different in achievement.

TABLE I

THE EFFECT OF EQUAL PRACTICE ON INDIVIDUAL DIFFERENCES WHEN MEASURED BY AMOUNT OF TIME PER UNIT OF WORK

	First Minute	Last Minute	Gain
A.....	5	15	10
B.....	12	30	18

Now let us measure efficiency by the amount of time per unit of work or examples;

TABLE II

THE EFFECT OF EQUAL PRACTICE ON INDIVIDUAL DIFFERENCES WHEN MEASURED BY AMOUNT OF TIME PER UNIT OF WORK

	Sec's. per Example, First Minute	Sec's per Example, Last Minute	Gain
A.....	12	4	8
B.....	5	2	3

Conclusions.—*A*, who had the lowest initial performance, gained much more than *B*. An equal amount of practice made them more alike. Hence, although individuals are born with unequal abilities, equal opportunity greatly reduces these differences.

Both conclusions are equally valid, and each one depends upon the method used for measuring improvement and individual differences, and both methods are equally valid for measuring efficiency. But where is the nigger in the woodpile? Let us take a more abstract case.

$$\text{Efficiency} = \frac{\text{work,}}{\text{time}} \text{ or } \mathrm{E} = \frac{\mathrm{W}}{\mathrm{T}}$$

Let *A*'s efficiency $= \frac{2W}{T}$ the first minute.

Let *B*'s efficiency $= \frac{W}{T}$ the first minute.

Let *A*'s gain $= 3G$.

Let *B*'s gain $= 2G$

Then *A*'s efficiency $= \frac{2W + 3G}{T}$ the last minute.

Then *B*'s efficiency $= \frac{W + 2G}{T}$ the last minute.

Now the argument from Table I. assumes that since $3G$ is greater than $2G$ therefore $(2W + 3G)/(W + 2G)$ is greater than $2W/W$. We know from elementary arithmetic that the value of a fraction is unchanged when both the numerator and the denominator are multiplied by the same number, but this is not true when equal or unequal numbers are added to the numerator and the denominator respectively. In interpreting the effect of practice upon improvement and individual differences we have been forgetting our denominators and have been assuming that we could compare fractions by simply considering numerators.

Let us then make this correction and study the effect of practice on the differences between A and B when we compare the ratios of their initial and final performances. From Table I.:

A's initial efficiency $= \frac{5}{60}$,

B's initial efficiency $= \frac{12}{60}$,

Ratio A to $B = \frac{5}{12}$;

A's final efficiency $= \frac{15}{60}$,

B's final efficiency $= \frac{30}{60}$,

Ratio A to $B = \frac{1}{2}$.

From Table II.

A's initial efficiency $= \frac{60}{12}$,

B's initial efficiency $= \frac{60}{5}$,

Ratio A to $B = \frac{5}{12}$;

A's final efficiency $= \frac{60}{4}$,

B's final efficiency $= \frac{60}{2}$,

Ratio A to $B = \frac{1}{2}$.

Now we can say that whereas A was 5/12 as efficient as B before practice he is one half as efficient after practice. That is, practice reduced the differences in the efficiencies between them. This method of measuring individual dif-

ferences gives the same interpretation for both methods of measuring efficiency and improvement. It is therefore justifiable to measure the effect of practice upon individual differences by comparing the ratio of the high to the low scores at the beginning of practice with those at the end of practice.

I shall now apply this measure to the experimental evidence cited in support of the quotations from Thornkike, Starch and Henmon given above.

Thorndike gives us three experiments of his own. I shall here reproduce portions of the tables from each.

TABLE III

1. THE EFFECT OF EQUAL AMOUNTS OF PRACTICE UPON INDIVIDUAL DIFFERENCES IN THE MENTAL MULTIPLICATION OF A THREE-PLACE BY A THREE-PLACE NUMBER

	Hours of Practice	Amount Done per Unit of Time		
		First 5 Examples	Last 5 Examples	Gain
Initial highest 5 individuals.....	5.1	85	147	61
Initial lowest 5 individuals but 1..	5.2	31	57	26

Comparing ratios we get:

Ratio of highest 5 to lowest 5 but 1, first 5 Ex's.....2.74
Ratio of highest 5 to lowest 5 but 1, last 5 Ex's.....2.58

Although the highest five gained much more than the lowest five, yet the difference between their efficiencies has been reduced.

TABLE IV

2. THE EFFECT OF EQUAL AMOUNTS OF PRACTICE UPON INDIVIDUAL DIFFERENCES IN COLUMN-ADDITION OF ONE-PLACE NUMBERS

	Average number of additions per 5 minutes corrected for errors.			Average Time Spent in Practice Minutes
	First Test	Last Test	Gain	
Initially highest 6 individuals..	297	437	140	40
Initially lowest 7 individuals...	167	345	111	46

Comparing ratios we get:

Ratio of highest 6 to lowest 7, first test.....1.78
Ratio of highest 6 to lowest 7, last test.....1.90

Here the difference in their efficiencies has been increased.

TABLE V

3. THE EFFECT OF 45 MINUTES PRACTICE IN WRITING THE PRODUCTS OF 11 x 11 TO 19 x 19 ON THE INDIVIDUAL DIFFERENCES OF 163 COLLEGE STUDENTS

	Approximate Output per Minute		Gain
	First Sheet	After 45 Minutes Practice	
Initially highest 6 individuals	15.4	24.2	8.8
Initially lowest 9 individuals	6.1	13.6	7.5

Comparing ratios we get:

Ratio of highest 6 to lowest 9, first sheet.....2.52
Ratio of highest 6 to lowest 9, end of 45 minutes.....1.78

Here the difference between the efficiencies of the two groups have been reduced in spite of the fact that the highest group made a greater absolute gain.

Starch gives us two experiments of his own in addition to citing evidence from other investigators.

TABLE VI

1. THE EFFECT OF MULTIPLYING 50 3-PLACE NUMBERS EACH DAY FOR FOURTEEN DAYS. EIGHT SUBJECTS

	No. of Examples in first 10 Min.	No. of Examples in Last 10 Min.	Gain in No. of Examples	Per Cent Gain
Three best persons	39	84	45	115
Three poorest persons	25	51	26	104

Comparing ratios we get:

Ratio of 3 best to 3 poorest first 10 min.....1.56
Ratio of 3 best to 3 poorest last 10 min.....1.64

Here the difference in the efficiencies between the best and the poorest has been increased.

TABLE VII

2. THE EFFECT ON INDIVIDUAL DIFFERENCES OF 120 MINUTES PRACTICE IN TRANSCRIBING LETTERS. 20 PERSONS

	First 5 Min.	Last 5 Min.	Gain
Ave., Initial highest five	139	310	171
Ave., Initial lowest five	100	239	131

Comparing Ratios we get:

Ratio of highest 5 to lowest 5, first 5 min	1.39
Ratio of highest 5 to lowest 5, last 5 min	1.26

Now let us average the ratios of the highest to the lowest scores before practice and those after practice in case of the 5 experiments cited.

	Ratio of Highest to Lowest before Practice	Ratio of Highest to Lowest after Practice
Thorndike 1	2.74	2.58
Thorndike 2	1.78	1.98
Thorndike 3	2.52	1.78
Starch 1	1.56	1.64
Starch 2	1.39	1.26
Ave's	1.99	1.85

Here we see that in two experiments equalizing practice increases individual differences but in three experiments it decreases them. In the first two 27 subjects were used and in the last three 212 subjects were used. The general tendency is then for equal practice to decrease individual differences.

The above experiments suffer from the fact that the practice periods were very short, the longest being 5.2 hours. At best we do not expect an education of that length to change anyone a great deal. We must therefore attach greater weight to Henmon's experiment since he measured the effect of a school year's work in arithmetic upon improvement from September to May in case of 47 pupils from grades 4, 5, and 6. Henmon ranks these pupils according to the score obtained in the Courtis tests in May and then compares the average absolute gains attained by each quartile in grade 4 and by each tertile in grades 5 and 6. I shall reproduce the September and May scores and the gains of the highest and lowest groups from each grade.

TABLE VIII

	Sept.	May	Gain
Average of 5 highest, Grade 4	8.6	33.2	24.6
Average of 5 lowest, Grade 4	1.4	10.0	8.6
Average of 4 highest, Grade 5	15.8	61.5	45.7
Average of 4 lowest, Grade 5	11.5	32.5	21.0
Average of 5 highest, Grade 6	24.6	63.0	38.4
Average of 5 lowest, Grade 6	12.3[1]	32.6	19.2

[1] Based on 3 pupils.

Comparing the ratios of the highest to the lowest in Septtember and in May we get:

TABLE IX

	Ratio of the Highest to the Lowest in September	Ratio of the Highest to the Lowest in May
Grade 4	6.14	3.32
Grade 5	1.38	1.92
Grade 6	2.00	1.93
Averages	3.17	2.39

Again we see that using the method of ranking given by Henmon, the general tendency is for the individual differences to be decreased by equal training.

It seems to us, however, that Henmon's method of ranking is undesirable. We are more interested in predicting the prospective scores from the September scores than we are in taking the May scores of the pupils and then finding what their scores were in the previous September. I therefore ranked the pupils according to their September scores and then calculated the average scores of the highest and lowest quartile in grade 4 and of the highest and lowest tertiles in grades 4 and 5.

TABLE X

THE EFFECT OF 9 MONTHS TRAINING ON ABSOLUTE AND RELATIVE GAINS OF HIGHEST AND LOWEST GROUPS IN ARITHMETIC WHEN PUPILS ARE RANKED ACCORDING TO SCORES AT BEGINNING INSTEAD OF END OF TRAINING

	Sept.	May	Amt. Gain	Per Cent. Gain
Average of initially 5 highest Grade 4	11	28.2	17.2	156
Average of initially 5 lowest Grade 4	.4	17.8	17.4	435
Average of initially 4 highest Grade 5	20.25	52.25	32	158
Average of initially 4 lowest Grade 5	7.25	30.50	23.25	320
Average of initially 3 highest Grade 6	30.66	63.33	32.67	106
Average of initially 3 lowest Grade 6	9.66	37.33	27.67	287

Comparing the ratios of the highest to the lowest scores for September we get:

TABLE XI

THE EFFECT OF 9 MONTHS TRAINING ON INDIVIDUAL DIFFERENCES WHEN MEASURED BY RATIOS OF HIGHEST TO LOWEST GROUPS AT BEGINNING AND END OF TRAINING. RANKS AS IN TABLE X

	Ratio of Highest to Lowest in September	Ratio of Highest to Lowest in May
Grade 4	27.50	1.58
Grade 5	2.78	1.71
Grade 6	3.17	1.69
Averages	11.15	1.66

Henmon interprets his results thus: "An inspection of the tables shows that by any rational method of interpretation the scriptural law 'to him that hath shall be given' holds. Those with the highest initial scores have the highest final scores and gain most absolutely and relatively." We now see how widely the statement deviates from the facts in the new order. Those with the highest initial scores do gain the most absolutely in two grades but not in all. When we compare the relative gains, we see that those with lowest initial score gain from 2 to 3 times as much as those with the highest initial scores. The result of the year's training is that the differences in the individual efficiencies are greatly reduced. In fact in September the highest groups were in the average over 11 times as efficient as the lowest group, but in the following May they were less than 2 times as efficient as the lowest. The difference in the ratio for grade three is extremely large, but this is just what we should expect when the initial efficiencies of some in the lowest group are zero. This is an encouraging fact to both pupils and teachers, for it means that we can accomplish something by training those who have not yet learned much.

Henmon also measures the effect of training on individual differences by comparing the deviations of the averages of the highest and lowest quartiles and tertiles from the average of the entire group. By this method he finds that the difference between the highest and lowest groups is much greater in May than it was in September. Murdock applied this same method to Henmon's results when the pupils were ranked according to their scores in September instead of May. She finds that when this is done the differences between the highest and lowest groups is less in May than it was in September. It is a poor rule that doesn't work both ways. We shall see a bit later that we can not measure the effect of training on individual differences by comparing the usual measures of variability before and after training. These results again depend on the measures of efficiency and improvement. If we use amount of work per unit of time, the distribution after practice is wider than before practice. If we use amount of

time per unit of work the distribution after practice is much narrower than before practice. But if we use the ratio of the highest quartile, octile or decile to the lowest quartile, octile, or decile before practice and after practice, then the effect of the training is the same no matter how we measure efficiency and improvement.

Another method that may be used to measure the effect of training on individual differences is correlation between initial performance and per cent of improvement. If the amount of improvement is proportional to the initial score then this correlation will be +1 or nearly so. This means that the individual differences are increased by training. But if those with high initial scores gain relatively less than those with low scores, then the correlation will be negative, or at the most, low positive; and we can say that the individual differences have been reduced by training. But we can not use correlation between initial and final performances, because it is quite possible for a bright individual to gain relatively less than a low individual without changing his rank in the first performance. A man with a million dollars capital gaining only 3 per cent. will still be a far richer man than one with a thousand dollars gaining 100 per cent.

Gates has calculated the correlations between initial performance and amount of improvement for reading and spelling in case of 105 pupils from Grades 3 to 8. Although he did not interpret his results at all from the standpoint of the effect of training on individual differences, yet we may use them in this connection. The first tests were given in November and the last in May. Comprehension was measured by the Thorndike-McCall Scale, and rate by a combination of the results from the Courtis, Monroe, and Brown's Scales. Spelling was measured by the Buckingham-Ayres Scale. The averages are given in Table XII.

Here the evidence seems clear that individual differences are reduced by training.

Further evidence in support of this conclusion is found in experiments which I now wish to report.

One hundred eight students practiced mirror drawing using a six pointed star. Each student did fifteen stars. The

time for each star was recorded in seconds. Improvement was measured by the times for the first and last stars.

TABLE XII[1]

The Effect of Equal Training on Individual Differences When Measured by Correlation between Initial Performance and Improvement	Correlation between Nov. scores and Amount of Improvement from Nov. to May
Comprehension of Reading	−.10
Rate of Reading	−.22
Spelling Grades 3, 4, 5, 6	−.67
Spelling Grades 7, 8	−.27

One hundred twenty-nine students practiced saying the alphabet backwards. Each student repeated the exercise 12 times. The experimenter made the subject halt to correct mistakes and recorded the time for each repetition. Improvement was measured by the times for the first and last repetitions.

One hundred seventeen students practiced an exercise substituting symbols for digits. Each student made 400 substitutions. In case of 58 students the experimenter recorded the time for every 20 substitutions. Improvement was measured by the times for the first and last 20 substitutions. In case of 59 students time was called every 30 seconds and improvement was measured by the number of substitutions for the first and last 30 seconds.

One hundred forty students practiced adding problems consisting of 5 2-place numbers having no zeros. Time was called every 10 minutes and improvement was measured by the number of examples correct in the first and last ten minutes.

To measure the effect of training on individual differences I calculated the following measures:

1. The q, or semi-interquartile range for the first practice and the last practice.
2. The total range between the highest and lowest scores of the first practice and the last practice.
3. The ratio of the highest quartile or q_3 to the lowest quartile or q_1, for the first practice and the last practice.

[1] A. O. Gates, 'Psychology of Reading and Spelling,' 1922.

4. The average absolute gain of the four highest and the four lowest for each exercise.
5. The relative gains for the same individuals.
6. The ratio of the average scores of the four highest to the four lowest for each exercise for the first practice and the last practice.
7. Correlations between the initial performance and the amount of improvement for each exercise.

Table XIII. gives the result for the first three measures, Table XIV. for the next three, and Table XV. for the last three.

Table XIII

The Effect of Equal Practice on Individual Differences in Skill in Mirror Drawing, Saying Alphabet Backwards, Substitution and Addition as Measured by q, Range, and Ratio of q_3 to q_1

	No. of Subjects	First Practice q	Last Practice q	First Practice Total Range	Last Practice Total Range	First Practice q_3/q_1	Last Practice q_3/q_1
Mirror Drawing	108	37.39	10.31	240	107	2.17	1.96
Saying Alphabet backwards	124	8.29	4.03	95	42	1.95	1.88
Substitution, Time per work unit.	58	12.61	5.80	95	25	1.63	1.75
Substitution, Work per time unit.	59	2.18	2.59	20	30	1.78	1.85
Addition	140	9.87	9.05	50	64	1.51	1.36
Averages of Ratios	.....	.....	.....	.....	.....	1.81	1.76

It will be seen that when efficiency is measured by amount of time per unit of work, the q's and ranges are greatly reduced by practice but when the efficiency is measured by amount of dork per unit of time, just the opposite is likely to be true. This is noticed particularly in the Substitution Experiment where both measures were used. We must therefore base our judgment upon the ratio of the upper to the lower quartiles before and after practice. Here it is seen that the tendency of equal practice is to reduce individual differences.

It will be seen that both the absolute and the relative gains are far greater for the 4 lowest than for the four highest and that when we compare the ratios of their initial and final performances the lowest and highest individuals are much closer together after practice than before practice. Whereas before

TABLE XIV

THE EFFECT OF EQUAL PRACTICE ON INDIVIDUAL DIFFERENCES IN SKILL IN MIRROR DRAWING, SAYING THE ALPHABET BACKWARDS, SUBSTITUTION, AND ADDITION AS MEASURED BY AVERAGE ABSOLUTE AND RELATIVE GAINS OF THE FOUR HIGHEST AND THE FOUR LOWEST AND BY RATIOS OF THE LOWEST TO THE HIGHEST BEFORE AND AFTER PRACTICE

	No. of Subjects	Amt. Gain Ave. of 4 Lowest	Amt. Gain Ave. of 4 Highest	Per Cent. Gain Ave. of 4 Lowest	Per Cent. Gain Ave. of 4 Highest	Ratio of the 4 Highest to 4 Lowest before Practice	Ratio of 4 Highest to 4 Lowest after Practice
Mirror Drawing	108	107.75	9.6	79.4	43.6	10.68	3.81
Saying Alphabet Backwards	129	72.5	9	73.1	34.6	14.60	5.03
Substitution, Time per Work Unit	58	73.8	1.75	70.3	1.07	3.92	1.24
Substitution, Work per Time Unit	59	3.07	2.00	49	1.03	3.01	2.26
Addition	140	15.75	10.00	64.9	14.70	2.82	1.93
Average of Ratio	.....	.....	.....	.....	.....	7.01	2.85

practice the four highest were in the average over 7 times as efficient as the lowest, after practice they are less than three times as efficient. The differences between the individuals at the extremes of the distribution curve are much more reduced than those near or within the semi-quartile range.

TABLE XV

THE EFFECT OF EQUAL PRACTICE ON INDIVIDUAL DIFFERENCES IN MIRROR DRAWING, SAYING THE ALPHABET BACKWARDS, SUBSTITUTION, AND ADDITION AS MEASURED BY CORRELATIONS BETWEEN INITIAL PERFORMANCES AND PER CENT OF IMPROVEMENT

	No. of *S*'s	*r*
Mirror Drawing	108	−.35
Saying Alphabet Backwards	129	−.41
Substitution		
Time per work unit	59	−.75
Work per time unit	58	+.02
Addition	140	−.74

Again we are obliged to draw the same conclusion: training reduces individual differences.

If we measure the differences in the efficiencies of individuals, by ratio of the highest to the lowest scores before and after practice then it seems clear that the general tendency is

for equal training to make individuals more equal in their achievement. This applies not only to the facts of our own experiments but also to the facts of investigators who have drawn an opposite conclusion.

Coming now to the general question of whether heredity or environment is the greater factor in causing individual differences it seems that we must be less enthusiastic about heredity than we have been. In the simpler skills at any rate, those that are taught in our elementary schools, and which are important for every one in making a living and in discharging the duties of citizenship in a democracy, the evidence indicates clearly that equal training makes individuals more equal in spite of their differences in native ability. Many and probably most of the inequalities of life could be reduced by giving every one equal opportunities in education. But when it comes to difficult problems requiring great native ability such as the mental multiplication of 3-place numbers, the solving of originals in geometry, the designing of machines, or the direction of great social movements, it is likely that here the achievments are more nearly proportional to native ability.

Summary

The conclusion of leading psychologists that equal training increases individual differences is erroneous because it is based on incorrect measures of individual differences.

Conclusions based on results when efficiency is measured by amount of time per unit of work are directly opposite those based on results when efficiency is measured by amount of work per unit of time. Therefore both measures are invalid for individual differences.

Reliable measures of the effect of training on individual differences are (1) the ratio of the high scores to the low scores at the beginning and at the end of practice, and (2) correlations between initial performance and percent of improvement.

When the first measure is applied to hitherto published results it shows that individual differences are decreased by training.

When both measures are applied to results obtained from experiments in Mirror Drawing, Saying Alphabet Backwards,

Substitution and Addition in each of which over 100 subjects were used, they show clearly that individual differences are decreased by training,

Equal training reduces more the differences between those at the extremes of the distribution curve than the differences between those nearer the median.

In the simpler skills equal education makes individuals more equal in spite of the differences in ability but this is probably not true for very difficult tasks.

A STUDY IN THE MEASUREMENT OF MUSCLE TONUS AND ITS RELATION TO FATIGUE [1]

BY ROLAND C. TRAVIS

The State University of Iowa

INTRODUCTORY

I. *Statement of the Problem.*—This experiment was undertaken to test the truth of the hypothesis, that various conditions which influence personal efficiency also affect postural tonus. If this hypothesis is true then the measurement of tonus should be an indirect measure of efficiency, a measure whose accuracy would, of course, depend upon the closeness of correlation between the two. If such a correlation could be shown to exist between postural tonus and efficiency one would have a valuable measure of the effect of such conditions as fatigue, daily ryhthm, drugs and such diseases as neurasthenia.

There is good a priori evidence that tonus might serve as a barometer of efficiency. The cause of tonus has been shown to lie in the nervous system rather than in the muscle itself; that is, it is neurogenic rather than myogenic; for if the nerve going to the muscle is cut the muscle loses its tone. It has also been shown that tone is a reflex phenomenon; for, if the afferent paths from the leg of a frog are destroyed without injuring the motor paths, the muscles lose their tone. Besides, tonus is affected by impulses from the higher centers as is shown by changes in tonus due to emotional and intellectual activity.

Tonus is a condition of readiness to respond and the success of the response is surely determined to some extent by this condition of readiness. This readiness is influenced by a vast number of factors some of which play an undetermined part. In a number of experiments it is almost impossible to control all the combinations which might affect the efficiency of the

[1] This paper reports an experimental study carried on in the Laboratory of the Psychological Clinic at the State University of Iowa under the direction of Professor John J. B. Morgan, Director of the Clinic.

subject. This being the case when the subject fails to respond to a change in certain conditions which have been controlled one is not able to state whether the controlled condition had no effect or not. Some undetermined factor might have compensated for the decrease in efficiency which the controlled factor would otherwise have produced. Since tonus is a sort of a composite result of all factors affecting readiness to respond some method of accurately measuring tonus ought to prove very desirable and valuable in the psychology of efficiency.

II. *Previous Work.*—There is very little psychological literature to be found bearing on muscle tonus as being a possible measure by which the mental accompaniments of fatigue and the factors affecting readiness to respond may be observed. Howell[1] made some experiments on fatigue and muscle tonus purely from the physiological standpoint. His observations go to show that the cause of muscle tonus is neurogenic rather than myogenic, that is, the cause lies in the central nervous system rather than in the muscle itself, and that there is very little energy consumed in the kind of contraction in the skeletal muscle designated under the term 'muscle tonus.' If the nerve is cut going to the muscle the tone is immediately lost, thus muscle tonus is concluded to be a reflex phenomenon and tonus may be excited reflexly through and by other sensory fields, *e.g.*, external cold by its action on the cutaneous nerve throws the muscles reflexly into a state of increased tonus.

The proper tonus is a very important factor in the skeletal muscles in maintaining the general posture of the trunk and the limbs, but tonus is of equal importance in the muscles of the heart, blood vessels, intestines and other viscera. Sherrington[2] asserts that in the spinal dog or cat the extensor muscles of the knee still possess considerable tonus, and that this spinal tonus is reflex, being traceable to the afferent nerves arising in these muscles themselves. Thus "the same exciting stimulus that reflexly throws the flexors into contraction interrupts reflexly the reflex tonus of the extensor muscles."

[1] Howell, 'Text-book in Physiology.'

[2] Sherrington, 'Integrative Action of the Nervous System.'

Ladd and Woodworth [1] found that the extirpation of the semicircular canals of one side resulted in a loss of compensatory movements due to a loss of tonus in the muscles of one side of the body. Head [2] observed that severe lesions of the cortex and subcortical tissues caused a complete atonic condition of the muscles. Tilney and Riley [3] observed that there was a type of cortical control, other than volitional control of the muscles, that gives rise to an 'inhibitory influence' on the motor cells of the spinal cord, and that this inhibitory control manifests itself in the tone of the muscles thus making possible the 'ideal muscle tone.' Hypertonicity is accompanied by an increase in the reflexes of the affected parts, and the conclusion is, that the motor cells of the cortex not only influence the muscle tone but tend to inhibit the action of the reflexes connected with those muscles.

Thorndike [4] has made some elaborate studies and experiments on the mental symptoms of fatigue. He brings out the fact that the many different kinds of work, such as, mental work, muscular work, glandular work, work of the red blood corpuscles, work of the enzymes, etc., make the classification of the different kinds of fatigue a very difficult matter. He classifies them under the following three heads: (1) Sensorial fatigue, which would include, for example, the diminished efficiency of the rods and cones in being set in action by light. (2) Intellectual fatigue, which is caused by the work done by the 'connection-system.' The reasoning and associative memory processes, etc., being the type of work done. (3) Muscular fatigue which is due to a diminished contraction of the muscles caused by a constant contraction of the muscles through work. He further states that, "if an individual continues in spite of his feelings of fatigue, the intensity of his feelings of fatigue is a very, very inadequate measure of his loss of efficiency. . . . The feelings of fatigue, from what little is known of them, thus seem a very poor symptom of the loss of ability."

[1] Ladd and Woodworth, 'Physiological Psychology.'
[2] Head, Henry, 'Studies in Neurology.'
[3] Tilney and Riley, 'Forms and Functions of the Central Nervous System.'
[4] Thorndike, E. L., 'Educational Psychology,' Vol. II.

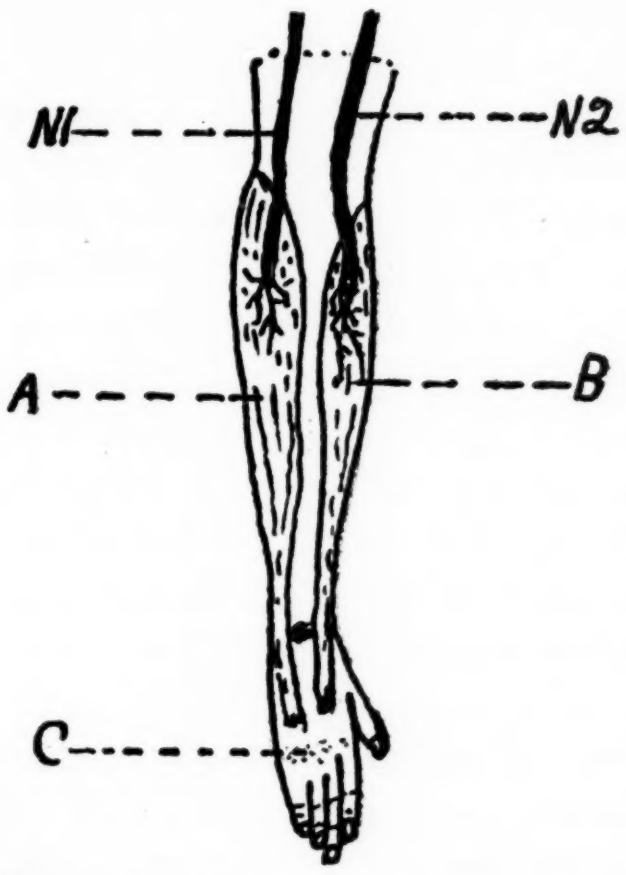
Most of the studies and experiments that have been made on this subject have been purely physiological, thus the need for a psychological study of the subject is apparent. Watson [1] asserts that the whole subject of muscular tonus is still a research problem.

III. *Method and Apparatus.*—The problem of approach to the measurement of muscle tonus has been rather difficult. Through trial, elimination and selection of materials an apparatus has been devised, the principle of which is correct, in spite of its miniature appearance. The apparatus used was devised by Professor John J. B. Morgan. It is an apparatus that measures the impact of a weight which swings at the end of an iron bar. This weight strikes the back of the hand which is held in a relaxed position on the proper place of the machine. The apparatus is composed of a swinging pendulum mounted in a frame 14 inches high, 14 inches long and 4 inches in depth. The pendulum is so constructed as to swing through 90 degrees. The weight at the end of the pendulum is composed of a rubber knob fastened to two sliding rods. When the rubber knob strikes the back of the hand it is forced back against the tension of the spring, the rods in turn force a pointer back along a scale. The distance this pointer is forced back depends upon the resistance offered against the kinetic energy of the swinging pendulum. There is an adjustible catch which allows the pendulum to be started from any point within the 90 degrees. The pendulum is caused to swing by its own weight and is initiated by bringing the pendulum above the catch, releasing the catch and allowing the pendulum to swing until the rubber knob strikes the back of the hand. The greater the tonus in the muscles of the arm the greater the reading on the scale will be. The accompaning diagram represents graphically how muscle tonus is measured.

We are assuming that *A* and *B* are in a state of relaxation, that *C* is in a position of rest, although free to move either to the left or right under the application of the required amount of external force. Now if *C* is caused to move either to the right or left, let us say for the sake of our illustration to

[1] Watson, John B., 'Psychology from the Standpoint of a Behaviorist.'

the left, by some external force, there will be a resistance set up in opposition to this move due to the normal state of tonus in muscle *B*. Thus, the tonus of *B* and *A* are equal in this normal state of relaxation, otherwise *C* would be moved either



A.....Extensor muscle.
B.....Flexor muscle.
C.....Lever through which extensor and flexor oppose each other.
*n*1.....Nerve to the extensor muscle.
*n*2.....Nerve to the flexor muscle.

to the right or left by the muscle receiving the greater stimulation from the central nervous system over nerve *n*1 or *n*2. So the muscle tonus of *B* is measured by the impact of the swinging pendulum against the back of the hand which is almost instantaneously forced back as a result of the impact. Now the greater the tonus in *B* the more resistance will be set up against the impact of the swinging weight, and the farther back the pointer will be forced along the scale thus causing a greater reading. This is the measurement of the initial impact rather than the reflex contraction which comes later. It is the measurement of the actual postural tension of the muscle.

The observer was seated comfortably before a table of average height upon which was placed the apparatus. The

observer was told to relax as much as possible, and especially to sit so that the arm and hand would be in a relaxed condition. A careful attempt was made to get each observer in the same position at each observation, particularly the right arm and hand which were used in these observations. It would be impossible of course to eliminate all variation in this respect, but the observations were carried on accurately enough to show a distinctive trend in results. An adjustable holder for the arm was used. It held the forearm rigid and permitted the hand to bend freely at the wrist. The arm was always in the same position, and the same leverage was applied to the hand regardless of the difference in length of the hands of the different observers. The holder is so designed as not to hinder the contraction and relaxation of the muscles. The experimenter sat at the opposite side of the table and the raising, the release and the swinging of the pendulum were all obstructed from the view of the observer by the flat board on the back of the apparatus.

The individuals acting as observers in this study were college juniors, seniors, or graduate students. Each observer was given two or three trials at the beginning of each observation to become accustomed to the striking of the pendulum on the back of his hand.

IV. *Results.*—Table I gives the results of an experiment carried on for the purpose of determining the difference between the normal muscle tonus of an individual and the condition of muscular contraction. It was attempted to put each subject esperimented on under the same external conditions as near as possible. Each subject was alone with the experimenter, and was told to relax as much as possible after his arm and hand had been arranged comfortably on the proper place of the apparatus in the arm holder. After each impact of the pendulum the arm and hand were returned to their original position and again relaxed. This was necessary for each impact of the pendulum slightly moved the arm and hand out of position, and in placing the hand and arm back in position necessitated a certain amount of tenseness in the arm muscles, thus every impact of the pendulum required a new

relaxation. There were 14 observers, and each observer was given five trials or impacts of the pendulum. The reading was taken after each impact, and the average of the reading of the five impacts was taken as the final reading which is represented in column 3 in Table I. for each observer.

TABLE I.

1	2	3	4	5
Observers	Sex	Normal tonus	State of Contraction	Percentage of Increase in tonus
A.....	M.	10	15.6	56
B.....	M.	10.3	19.7	91
C.....	M.	9.7	16.3	68
D.....	F.	8.3	14.2	71
E.....	M.	9.8	14.	42
F.....	M.	8.8	19.4	120
G.....	M.	9.5	16.	68
H.....	M.	7.2	15.4	113
I.....	F.	6.6	16.2	145
J.....	M.	12.1	17.4	43
K.....	M.	9.8	17.	73
L.....	M.	7.4	16.	110
M.....	M.	7.8	16.2	107
Average.....		9.65	17.1	85

The observer was now told to clinch his fist so as to bring the muscles of the arm and hand into a state of contraction, and the observation was carried out as before, that is, each observer was given five impacts of the pendulum and the average of the five readings taken as the final result which is given in column 4 in table 1 for each observer. The latter observations were made at the same sitting as those given in column 3. The time taken for each observation was not more than ten minutes.

A comparison of columns 3 and 4 will show a vast difference between the relaxed state or the state of normal tonus and the state of muscular contraction. It ranges all the way from a 42 per cent. increase in observer *E* to an increase of 145 per cent. in observer *I*. It will be observed that two of the observers were women and the rest men. The adult average, regardless of sex, shows an increase of 85 per cent. in the muscular contraction state over that of the normal relaxed state. It will be noted that there is quite a wide variation be-

tween individuals both in the relaxed state and in the tense state. Observer *I* who is a woman ranks the lowest in the relaxed state and yet ranks relatively high in the state of contraction.

Table II. represents the results obtained from observations made for the purpose of determining the effect of muscular activity upon muscle tonus, and the difference between a normal state of tonus on arising in the morning and the normal state of tonus on retiring in the evening. Would the tonus of the skeletal musculature be greater or less after a short period of rather strenuous muscular exercise? And, would the tonus be greater in the morning on arising or greater in the evening on retiring?

TABLE II

1	2	3	4	5	6	7	8
Observations	Observers	Date	Tonus on Arising	Physical Exercise			Tonus on Retiring
				Tonus before	Tonus after	Time	
1	*A*	Feb. 7		12	11.4	3:30–5:30 P.M.	
2	*A*	" 9	10.1				11.5
3	*A*	" 10	10.1	11.3	9.5	4–6 P.M.	10.
4	*A*	" 14	9.4				10.8
5	*A*	" 15	9.7				10.4
6	*A*	" 16	9.9				11.
7	*A*	" 18	9.				9.7
8	*A*	" 19	9.	9.5	8.9	4–6 P.M.	9.8
9	*A*	" 21		10.4	10.	3:30–5 P.M.	
10	*B*	" 7		11.7	11.1	3:30–5:30 P.M.	
11	*B*	" 8	9.7				11.3
12	*B*	" 9		10.7	10.5	3–5:30 P.M.	
13	*B*	" 10	10.1	10.9	9.6	4–6 P.M.	10.3
14	*B*	" 13	10.3				11.4
15	*B*	" 14	9.3				10.9
16	*B*	" 15	9.				10.8
17	*B*	" 16	9.6	9.6	9.2	4–6 P.M.	10.2
18	*B*	" 18	9.				9.4
19	*B*	" 19	9.2	10.	10.	4–5:30 P.M.	10.2
20	*B*	" 21		8.8	7.5	3:30–5:30 P.M.	
A's Average....			9.6	10.8	9.95		10.45
B's Average....			9.52	10.28	9.65		10.56

Observers *A* and *B* acted as subjects in these experiments, being chosen not for any particular trait or characteristic, but

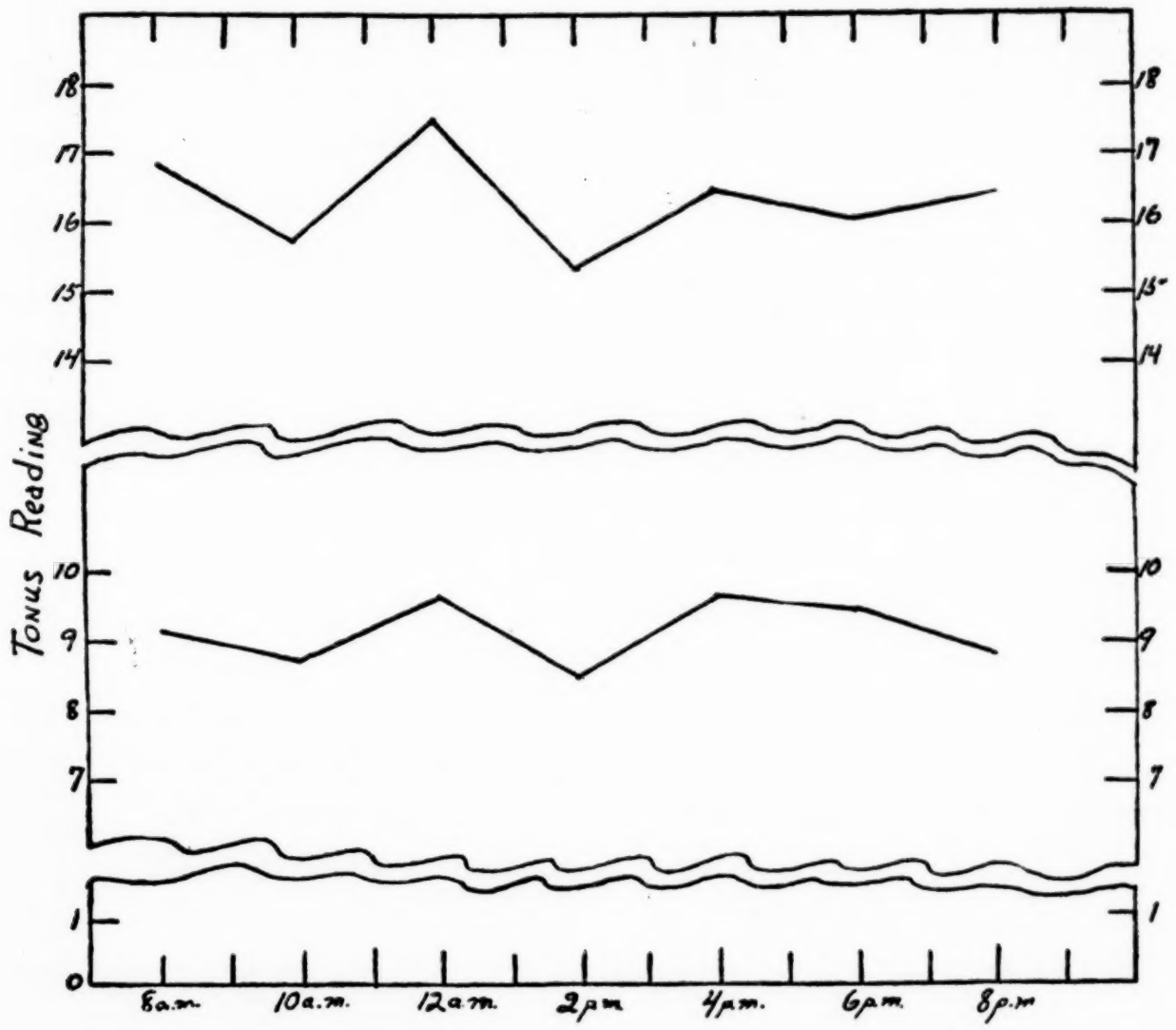
because they were in the habit of taking a short period of exercise each day. Each observer's tonus was measured in the normal relaxed state immediately before exercise and again shortly after exercise. The observer was told to relax as much as possible each time, and it was attempted to get the external conditions of the experiment the same after exercise as they were before exercise as to the bodily posture, the position of the arm and hand, and the mental set towards the relaxation. The apparatus was always in the same position on the table.

Column 1 in Table II. represents the number of observations regardless of the observers. Column 2 indicates the observers, and column 3 gives the dates of the observations, thus the observations continued over a period of two weeks. Columns 4 and 8 will be referred to later. Column 5 gives the normal tonus reading immediately before exercise, and column 6 the normal tonus reading shortly after exercise. Each reading represents the average of five readings of the impact. It will be seen that in all cases with the exception of observation nineteen that the normal tonus lowered as a result of the exercise. Apparently, *B*'s tonus was not changed that day as a result of the exercise.

The figures in column 4 in Table II. represent the normal tonus on arising in the morning and the figures in column 8 represent the normal tonus on retiring in the evening. These observations were carried on similar to the previous ones, thus being spread over a period of several days. Several of these observations were made during the same days as those for determining the effect of muscular exercise on muscle tonus The results indicate a greater tonus on retiring than on arising. This is true in every case with the exception of observation three, the tonus being 10.1 on arising, 11.3 at 4 P.M. just before exercise, 9.5 at 6 P.M. right after exercise, and 10 on retiring, thus the tonus of observer *A* was lower on retiring than on arising, and it was lower right after excerise than any other observed time of the day. It will be seen that there is some variation in the tonus of the same individual from day to day. In observations eight, thirteen, se-

venteen and nineteen the tonus lowered during exercise, then after several hours of rest went higher than any observed time during the day.

The graph represents the results of an experiment made on observer *B*. The normal tonus was measured every two hours throughout the day, from 8 A.M. to 8 P.M. The observer took no exercise that day and on the whole spent a day of leisure. It will be seen that at 8 P.M. the tonus was lower



than any other time of the day. This result is quite the reverse of the other observations. The lower curve in the graph represents the normal relaxed state or the normal tonus; the upper curve represents the state of contraction.

Summary of Results

1. There is a great comparative increase in the state of muscular contraction over that of the normal state of tonus. The average of 13 cases showing an increase of 85 per cent.

There is a wide variation between the different tonic states of different individuals.

2. Physical exercise decreases muscular tonus, that is, the elements entering to cause physical fatigue also react to cause a hypotonus in the skeletal musculature. This is probably purely a physical reaction. In some cases there was a considerable degree of 'feelings of fatigue.'

3. Just what would cause the rise in muscle tonus at the time of retiring over that of arising is a question perhaps not so easily answered. Whether it is due to a certain mental set, the nature of responses during the day, the physiological reactions during the day or a combination of the three is difficult to determine. But in the most of these observations the rise in muscle tonus at the time of retiring was accompanied by a 'feeling of fatigue.' This may suggest that perhaps the fatigue is more of a mental attitude rather than a physical condition.

4. The variation in postural tonus from day to day in the same individual is quite marked. There is considerable variation throughout the day in an individuals' postural tonus as shown by the graph.

Future Research

The studies of muscle tonus as affected by muscular and mental fatigue offer another more or less restricted field of research pointing toward a possible better interpretation of the mental and physical symptoms of fatigue, and perhaps the finding of a more definite clue into the so-called 'feelings of fatigue.'

It would be of interest to make some observations on common laborers by measuring their tonus before and after a hard day of physical labor and compare the results with the results obtained from the readings of the tonus of an intellectual student before and after a hard day of intellectual labor. Would the results show a comparison similar to the comparison of columns 5 and 6 to those of 4 and 8 in Table II, that is, columns 4 and 8 representing the tonus before and after of the student; and the columns 5 and 6 representing the tonus before and after of the laborer.

It would be of interest to measure the tonus of an individual while in the state of abstraction, while asleep, while under the influence of hypnosis, and compare results with the normal tonic conditions.

Would it be possible to detect by the tonus of the different muscles of the musculature any mal-functions of the motor areas of the brain?

Could an individual with slow muscular movements and long endurance be detected from the opposite type by the measurement of muscle tonus?

This elementary study may suggest the possibility of devising a test to determine whether a laborer in industry or a patient in the clinic really was in a condition of physical fatigue or whether the so-called 'feelings of fatigue' were due to a certain mental attitude—perhaps a lack of interest in his work.

EFFECTS OF MUSIC UPON ELECTROCARDIOGRAMS AND BLOOD PRESSURE

BY IDA H. HYDE

State University of Kansas

Received for publication July 19, 1923.

A scientific employment of the power exerted by music for specific purposes, as for instance to lessen nervous tension or fatigue, or to arouse emotions requires not only a knowledge of the listener's preference for certain selections of music or for a special musical instrument, but it is also essential to know the psychological as well as the concomitant physiological reactions that are produced by the music. These reactions as is well known are the result of nerve stimuli on definite tissue cells and can be demonstrated with the aid of suitable apparatus.

The purpose of the present investigation was to ascertain the effects of different kinds of musical selections upon the cardiovascular system in individuals who are known to be fond of music, persons indifferent or not sensitive to music, some Indian students, neurasthenics and some animals. The plan was to compare the effects of vocal music and that of different kinds of instruments upon listeners of different native endowment and training and under varying conditions, but only a beginning was made, enough however to prove that investigations in this field offer important returns.

The preliminary experiments were conducted on young men fond of music and under fairly constant subjective and weather conditions. From these first experiments (No. 1) it was learned that contrasted selections of music produced similar responses in pulse rate, pulse and blood pressure, velocity of the blood flow and action current, or electrical phenomena of the cardiac muscles. The data obtained from the selections that were familiar to the young men indicated that associated memory played a part in the results, and that the same selections produced a different effect when sung than when played

on the violin. It was interesting to see that meals, certain foods, smoking, vitiated air, fatigue, familiarity, and repetition, and excited and depressed states of mind all exerted an influence upon the effects produced by the music, and these varied for different individuals.

The same method that had been adopted in the preliminary investigation was employed in the continuation of the research. Of the fifteen men and women selected as listeners four were Indian students, seven were male students, two of which were not fond of music and could not distinguish one tone from another; and eight were women. Of the latter two were not sensitive to music, one was hysterical, one was an instructor in music, one had a defective heart valve and the others were fond of music.

The items included in each record were pulse rate, systolic, diastolic, and pulse pressures, measured in millimeters of mercury, relative velocity of the blood flow, and electrocardiograms that record the action current or electromotive force of the ventricular muscle contractions. These items were secured from one to five minutes before the music began, and from one to fifteen minutes after it had ceased. Until the persons were accustomed to the method of procedure, while listening to the music, only the cardiograms were taken, because it was found that the latter were affected by the manipulations necessary to secure the blood pressure records.

The electro-motor changes in the contraction of the cardiac muscles were obtained with an Einthoven string galvanometer. Through it, we have as is well known, for the first time a method for measuring and contrasting cardiac excitations and indirectly also changes in pulse rate and an aid in explaining changes in blood pressure.

For the comparison of musical and other stimuli the galvanometer was set for a film speed of 2.5 centimeters per second, and the time marker deflection per one millivolt. Lead 'Two' was adopted for all the curves, that is, when the persons right arm and left leg are connected to the galvanometer through silver cuffs kept in place by bandages moistened with physiological salt solution.

The photographed curves of the electrical currents show the changes in excitation of the heart muscle, and afford a novel method of approach to the study of changes due to pathological, psychological, or physiological influences. Three photographs were taken on the effect of each of the musical records from each of the listeners. These photographs were compared, tabulated, and interpreted for final conclusions.

A Tycos sphygmomanometer, controlled by an Erlanger's type of the apparatus was used to get pulse rate and systolic and diastolic blood pressure. From these records the pulse pressure and the relative velocity of the blood flow were calculated. (No. 2) The systolic pressure was read at the onset of the first phase and the diastolic at the beginning of the fourth in every case.

The following three records that seemed to be free from associated disturbing influences were decided upon for the comparison tests for all the persons. First the record of the Boston Symphony Orchestra of Tschaickowsky's Tragic Symphony characterized by slow minor movements. Second, the Toreador's brilliant description of the bull fight, from Carmen, as sung by Amato, and third the National Emblem, a stirring rhythmical march by Sousa's Band. An Indian war dance sung or rather whooped to an accompaniment on an Indian drum by the composer of the music, Prof. C. S. Skilton, was tested on most of the students. On a few special cases a lullaby played on the violin was employed to test the soothing restful effects upon the cardio-vascular system.

The individuals were classified in two groups. In one were placed those that had love for and training in music, and in the other those that lacked these. When similar results were obtained from several members of the same group, the average data only were tabulated, but unusual changes were noticed and taken into consideration in the final reports.

For control data the student's records were taken without the influence of music at different hours of the day but not directly before meals or two hours after meals because it was seen that ingestion of food, its quantity and character, affects the cardio-vascular records.

It was also found that cardio-vascular changes vary somewhat during the day, and for some individuals and for certain hours more so than for others. These variations may be related to the general diurnal metabolic and physical activities of the organism and are taken into account in preparing the data. As a rule the pulse rate and relative velocity of blood flow decreased at the end of an hour's experimentation. The general results of this set of experiments agree on the whole with those published by Weysse and Lutz No. 3. The results of the experiments that were conducted in a poorly ventilated room or upon students after smoking cigarettes, or on a convalescent patient, differ in certain particulars from those obtained from normal students under normal conditions without the influence of music. With poor ventilation systolic and diastolic pressure increased and other records decreased. Smoking of long duration lowered the pulse rate and pulse pressure, and relative velocity of the blood flow. But the hour's rest while sitting for the test proved beneficial to the convalescent person whose systolic and pulse pressure and relative velocity of the blood flow all increased.

Three tests were made for each of the thirteen listeners on the influence of the orchestral music of Tschaickowsky's Symphony. Among these were included individuals of different degrees of native endowment, training and physical conditions. The results obtained from these as demonstrated in the cardiograms vary in certain respects. Curves 60 and 62 recording the reactions from a male music student were selected as typical for the curves of the normal listeners fond of music. Curve 60 was recorded before the symphony was heard and 62 seven minutes after it had been listened to. In comparing these curves it was seen that QR, the highest deflection and the one that records the action current of the ventricular muscle contraction covers seven scale divisions of two-tenths of a millivolt each, and equal therefore to an electro-motive force of one and four-tenths millivolts, and the pulse rate had fallen to seventy-six per minute; A decrease of two-tenths millivolts and four heart beats per minute, due to the influence of the symphony was thus recorded. According

to the statements of the listeners the minor tones of the symphony aroused depressing sensations, these however were not of the same degree in all persons.

In considering therefore the information gained from a study of electrocardiograms, together with the average results recorded for all of the normal students before the symphony was heard with the data recorded from one to seventy minutes after it had ceased, we learned that on the whole, there was a decrease in function of the cardio-vascular system excepting in the diastolic blood pressure, by the reflex stimulation of the mournful tones of the symphony.

Due to a lack of sufficient data it is not possible to state at present how long the after effects of the minor tones persist. The effects may however be quickly counteracted by certain selections of music that inhibit the depressed condition. This was demonstrated by presenting fifteen minutes after the symphony had ceased the National Emblem March as played by Sousa's Band. Other cheering stimuli may have the same effect. On the other hand, the listener is placed by the minor tones in a receptive or susceptible mood for stimuli of a discouraging or depressing character.

For the purpose of testing the effect of the Symphony on a convalescent person responses were secured from a hysterical patient. Cardiograms were taken before, and one minute after, and also ten minutes after the convalescent had listened to the music. It was apparent that the minor tones affected and distressed her. Within a minute after hearing the symphony her systolic pressure rose from 112 to 118 mm. Hg., diastolic from 68 to 80, pulse rate from 78 to 90 per minute, but E.M.F. fell from 0.8 to 0.65 and the pulse pressure from 44 to 38. But ten minutes later the patient seemed faint and it was surprising to see that the records were reversed, showing a fall below those obtained before the music was heard.

Another interesting experiment was the influence of the symphony on the instructor in music, fatigued after a day's teaching. After listening to the music she remarked that she disliked it. The curves and data showed that five minutes after the music ceased the cardio-vascular records excepting

the pulse pressure and E.M.F. of the heart muscle increased and then fell. In both of the last two types of individuals there was, after hearing the music a marked rise of all the cardio-vascular functions excepting the pulse pressure and E.M.F. of the heart muscle. This was followed ten minutes later by a remarkable reversed action of the cardio-vascular activities which in the case of the convalescent was a change to less than what they were before the symphony had been heard.

As a result of these experiments it is safe to say, that music of the character of the symphony is not to be recommended for individuals who are fatigued, depressed or ill. It might be employed to subdue hilarity in individuals or masses of people.

The average results on the influences of the symphony on persons neither fond of nor sensitive to music were next tabulated and electrocardiograms obtained before the symphony had been heard and after it had been repeated several times. In comparing the cardiograms and data furnished by this set of students, it was seen, that the results obtained before they listened to the music, with those secured after they had heard it were practically alike, and very different from those of Group *A*. Evidently for the persons that lacked sympathy for music the minor symphony tones neither inhibited nor stimulated the nervous control of the functions of the cardio-vascular system under consideration, nor the emotions of depression or sadness.

The influence of Toreador's song in Carmen, describing the bull fight was next investigated. The cardiograms recorded from the listeners that were fond of and sensitive to music demonstrated that the song did not have the same physiological or psychological effect on all of the listeners. In those persons that had enjoyed the song very much, all of the reactions excepting that of the diastolic blood pressure were augmented. But three of the listeners did not enjoy the song, and their reactions were not increased but more or less lowered. The Indian student remarked that there was a challenge in the spirit of the song that annoyed him. It was found that his reaction had not become greater. The instructor of music

was familiar with the song, had very often heard it and did not especially care for it. In her case the results showed that the effect of familiarity and repetition of music that was indifferent had little or no effect on the cardio-vascular system. The song disturbed a listener who was in poor health, but a month later when she was quite well she enjoyed the song and the reactions due to it were then reversed, that is increased. A study of the cardiograms and records secured from persons not fond of or sensitive to music showed that the Toreador's song exerted neither a psychological nor a physiological influence on these persons. As a result of these experiments we may conclude that the brilliant tones of the Toreador's song stimulated the cardio-vascular functions to increased activity in those individuals that found pleasure in the song. On the other hand it did not augment the functions in those that did not enjoy it. Moreover the song exerted no influence on the persons that are not fond of music. It is not known from these experiments whether it was the spirit inherent in the song or the musical and vocal tones that produced the greater effect.

We shall now consider the influences of the rhythmical National Emblem played by Sousa's Band. On all of the sensitive listeners excepting on the music teacher, that did not care for this sort of music, the rhythmical Band music had a stimulating effect, and proved enjoyable. Moreover it was especially the systolic and pulse pressures and relative velocity of the blood flow that were stimulated to increased activity by the stirring tones. It was interesting to see how quickly this music had a bracing effect and removed fatigue. Then too the depressed responses to the minor tones of the tragic symphony were counteracted and again restored to their normal activities by the rhythmical sounds of this gay composition rendered by the musical instruments. On two of the non-sensitive listeners who recognized a difference between this and the other selections, this march produced a slight and increased reaction. But neither the stolid Indian girl nor the man who could not keep step seemed to be affected by even this class of music or instruments.

It finally became of interest to ascertain the effects of an Indian war song that was whooped and sung by the composer of the song to the drum accompaniment. The listeners were unprepared for this performance. It was not surprising therefore that the weird war tones and beating of the drum affected each of the listeners differently. On a male music student much interested in the composition of sounds it first produced a great rise in systolic and pulse pressure, 106 to 114 and 56 to 62 respectively, but a fall in all of the other reactions for about fifteen minutes duration, during which the pulse had fallen from 84 to 75 per minute. But in a convalescent, the tremendous effect was at once a marked decrease, especially of the systolic pressure 110 to 98, and velocity of the blood flow that lasted more than thirteen minutes. Judging from impressions and records obtained during and after the performances the effect on her was more like a shock. Also on a woman fond of music who had never heard anything of the kind, the sounds produced a shock like-effect resulting in a fall of all of the cardiovascular activities, excepting that a most remarkable increase in the electromotive force of the cardiac muscles took place.

On the other hand on an Indian man fond of music and familiar with Indian songs, the sounds produced surprise and pleasure. His records during the time of the performance showed an increase in 10 beats in pulse rate, 6 mm. in diastolic blood pressure and 0.2 millivolts in electromotive force, thus showing a striking difference in results from those above noted. The performance had some effect on an instructor in music who had often heard the song and was always amused by it. Here then was a type of music that actually had a psychological and physiological effect on the non-sensitive Indian listeners. As soon as the wild war song and sounds of the drum fell upon their ears the man, belonging to an Indian football team, was agreeably surprised. All his records excepting the pulse rate greatly increased during the rendering of the song and for at least eight minutes afterward. It seemed however that the unexpected performance suddenly robbed the Indian woman of her stolidity and left her in a sort of dazed condition. Her reactions excepting the diastolic pressure fell at once and

remained below normal for about eight minutes. It may be said however, that on the whole more often as a first result, the rendering of this selection produced a depression of the cardio-vascular activities on those listeners who had been unaccustomed to this kind of performance or for whom it recalled emotional associations, or who found it distressing. On the other hand for those Indian listeners who enjoyed it and for whom the surprise was not powerful enough to counteract the pleasure of the song, the reactions of the cardio-vascular system under discussion were increased. The effect of this wild war song may be likened to a reflex shock produced by strong stimulations. The after effect or emotional durability varied for the different individuals.

Another experiment was undertaken to test the effect of a lullaby played on a violin, upon a woman patient who appeared to suffer from nutritional disturbances as a consequence of influenza. The electro-cardiac records however revealed a disorder of the heart's action; an extreme auricular acceleration, or auricular flutter. The auricular contractions were conspicuously displayed by the shadow of the needle on the camera, and the influence of the lullaby on the patient distinctly observed in all its details. Both her husband and the experimenter were amazed at the suddenness of the change in the cardiac contractions when the lullaby was heard. In some of the tests the deflections of the flutter decreased, in others they apparently ceased. The systolic and pulse pressure, and relative velocity of the blood flow were increased, but the pulse rate and the electromotive force of the ventricular muscle decreased. The musical tones proved restful and beneficial to the patient. From the conspicuously beneficial effects produced upon the activities of the heart and the tonicity of the cardiovascular system in general, it is safe to recommend the lullaby as a sedative influence for individuals who are sensitive to musical tones.

The selections employed in this investigation may not be considered ideal ones for testing the effects of music upon individuals whose native musical endowment and training are very different. Nevertheless it has been demonstrated

through their use that certain selections of music and most likely certain musical instruments and qualities of the vocal sounds exert a far reaching influence upon the cardio-vascular system and very likely upon other functions of the body. It is probable that the employment of certain selections of music will prove a valuable aid in the treatment of nervous disorders.

To those persons who kindly gave their time to aid in this investigation I take this opportunity to express my deep appreciation.

Summary

It has been discovered that cardio-vascular functions are reflexly stimulated concomitantly with psychological effects of music and that, through the use of the Einthoven string galvanometer, and sensitive sphygmomanometers, the physiological reactions that have been excited by different sorts of music can be measured and compared.

Moreover with this method the proper sort of music may be prescribed and thus a scientific employment of the power inherent in music may prove a valuable adjunct to psychotherapy in the treatment of convalescent or other patients sensitive to music.

On persons not susceptible to music the tragic minor tones that characterized Tschaickowsky's symphony were without effect. But in persons endowed with musical sensitivity the tones of the selections produced a stimulation that as a rule reflexly lowered the functions especially considered in this investigation. This class of music is therefore not to be recommended for individuals depressed, fatigued or convalescent, but may be employed to subdue hilarity. The general effect of the Toreador's song was an increase in the functions of the cardio-vascular system. The results varied with the individual and depended upon the health, native endowment, musical training, also upon the interpretation of the theme and familiarity with the selection, and moreover upon the degree to which it was enjoyed or disliked. The song exerted no influence upon those listeners that were not sensitive to music or those that were familiar with the song but did not care for it.

As a rule, the records secured on the effect of the National Emblem March, showed an increase in the cardio-vascular activity, especially noticeable in the velocity of the blood flow and systolic and diastolic blood pressure. But for those who were not able to keep step with the march and lacked fondness for music the records remained unchanged. This sort of music proved valuable in counteracting depressed reactions produced by the minor tones of the symphony, in removing fatigue, and arousing muscular activity.

The song of the Indian war dance accompanied by the drum was the only music that exerted an influence upon the stolid non-sensitive Indian listeners. The Indian man was agreeably surprised by the performance. His records rose far above his normal ones. The Indian woman however was bewildered by the unexpected thrilling sounds and all of her reactions excepting the diastolic blood pressure were greatly decreased. The psychological and concomitant physiological reactions, excited partly by surprise and partly by the startling barbaric combination of tones varied for the different sensitive listeners.

From the conspicuously beneficial results exerted upon the activities of the heart, inhibiting auricular flutter in a patient and increasing the cardio-vascular tonicity in general, it is safe to recommend a lullaby played on the violin as a sedative for all individuals who are sensitive to musical tones.

We may conclude from the results of this investigation that most people are unfavorably affected psychologically and physiologically by music that is characterized by tragic mournful tones, and favorably affected by gay rhythmical rich toned harmonic melodies. Individual differences in native endowment and training are accompanied by individual differences in physiological reactions to certain musical compositions.

The indications are, that those selections of music rendered either vocally or instrumentally that exert a favorable reflex-action on the cardio-vascular system, have also a favorable influence upon the muscle tone, working power, digestion, secretions and other functions of the body.

Vocal and instrumental music may be selected that will excite psychological and concomitant cardio-vascular reactions the effect of which might inhibit irritability, act as a sedative, arouse optimism and be used as a valuable agent to scientifically organized labor.

References

1. Hyde and Scalapino. *Am. Jr. Physiol.*, 1918, XLVI.
2. Erlanger and Hooker. Johns Hopkins Hospital Rep., 1904, XII., 145.
3. Weysse and Lutz. *Am. Jr. Physiol.*, 1915, XXXVII., 347.

THE INFLUENCE OF THE CHANCE IMPERFECTIONS OF MEASURES UPON THE RELATION OF INITIAL SCORE TO GAIN OR LOSS *

BY EDWARD L. THORNDIKE

Institute of Educational Research, Teachers College, Columbia University, New York

When the individuals in a varying group are measured twice in respect to any ability by an imperfect measure (that is, one whose self-correlation is below 1.00), the average difference between the two obtained scores will equal the average difference between the true scores that would have been obtained by perfect measures; but for any individual the difference between the two obtained scores will be affected by the error. Individuals who are below the mean of the group in the first measurement will tend by the error to be less far below it in the second, and individuals who are above the mean of the group in the first measurement will tend by the error to be less far above it in the second. The lower the self-correlation, the greater the error and its effect.

All our measures of growth, practice, fatigue, or the effect of any cause producing difference in score in a second trial, are complicated by this effect, if we try to put them into relation to the amount of initial ability or eventual ability; or to compare groups not identical in initial ability; or to make up groups by pairing so as to have groups alike in initial ability.

The fundamental theory of the matter is the same as leads to the correction of correlations for attenuation [Spearman, '04], and to the correction of variabilities and measures of overlapping [Thorndike, '13, p. 93, and Kelley, '19]. The effect of imperfect measures on pairing for initial ability has been noted by Kelley, but the effect upon correlations of initial ability with gain or loss seems not to have been noticed. So it seems desirable to emphasize the general theory by illustrations, to state the formula for correction in one impor-

* This investigation is one of a number made with the aid of a grant from the Carnegie Corporation.

tant case, and to note some of the errors and inadequacies due to neglect of this effect of the error of measurement.

If we take a case where there is really no gain at all, *e.g.*,

Individual	True 1st Score	True 2d Score
a	5	5
b	6	6
c	6	6
d	7	7

and apply a chance error to it, we have a very clear illustration of the nature and amount of the error. I thus took 100 cases, one 1, 1, three 2, 2's, six 3, 3's, twelve 4, 4's, eighteen 5, 5's, twenty 6, 6's, eighteen 7, 7's twelve 8, 8's, six 9, 9's, three 10, 10's, and one 11, 11, and applied at random an error of the magnitude: − 4, one; − 3, four; − 2, eleven; − 1, twenty; 0, twenty-eight; + 1, twenty; + 2, eleven; + 3, four; + 4, one, to obtain the imperfect first-trial score. Again in the same way, I obtained an imperfect second-trial score. The result is Table 1, with a correlation of .5½ between the two measures and with spurious gains and losses running, in relation to the initial scores, + 1.33, + 1.20, + 1.67, + 1.06, + .83, − .17, − .64, − 1.67, − .55, − 2.8, − 2.33. If there had been a real gain of 2 points for all, the error would have made it appear to vary after the fashion, + 3.33, + 3.20, + 3.67, + 3.06, + 2.83, + 1.83, + 1.36, + .33, + 1.45, − .80, −.33, the slope of the relation line being the same as before.

TABLE I

THE EFFECT OF APPLYING A CHANCE ERROR TO TWO SERIES OF MEASURES WHICH IN THEIR ORIGINAL FORM SHOWED ZERO DIFFERENCES

Measure II. \ Measure I.	1	2	3	4	5	6	7	8	9	10	11
1	1	1									
2	1			2	1		1	1			
3		2	2	2	1	1		1			
4	1	1	1	1		3	1			1	
5		1	1	4	2	4	2	1	1		1
6			1	4	3	4	1	2		2	
7			1	4	4	4	2		1		
8						1	3	2	3		
9					1			1	3		
10						1	1	1	3	2	1
11											1

Measure I.

Measure II—Measure I.	1	2	3	4	5	6	7	8	9	10	11
−6								1		1	1
−5							1	1			
−4									1	2	
−3					1	1	1	1			
−2				2	1	3	2	2	1		
−1		1		2		4	1		3		1
0	1		2	1	2	4	2	2	3	2	1
1	1	2	1	4	3	4	3	1	3		
2		1	1	4	4	1		1			
3	1	1	1	4			1				
4			1		1	1					
5											
6											
n	3	5	6	17	12	18	11	9	11	5	3
Average gain	1.33	1.20	1.67	1.06	0.83	−0.17	−0.64	−1.67	−.55	−2.8	−2.33

Consider now the correlation between initial ability and gain if we should take these imperfect measures at their face value. It would be markedly negative (− .5 in our illustration), although in reality there is zero correlation between true first score and true gain.

In general, imperfection in the original measures will distort the relation between true initial ability and true gain, making it more negative. The relation between true final ability and true gain will be distorted in the opposite direction. Tables II. to VI. give illustrations. For example, let the true initial abilities be: one 1, three 2's, six 3's, twelve 4's, eighteen 5's, twenty 6's, eighteen 7's, twelve 8's, six 9's, three 10's, and one 11. Let the true relations between initial score and final score and between initial score and gain be as in Table II with a clear positive relation (about .6½) between initial ability and gain. Apply the same chance error as before, and the relation between initial obtained imperfect score and gain as computed from the imperfect measures is as shown in Table III. The real positive relation is reduced to below zero (about −.1½).

Suppose, on the other hand, that the relation of true initial ability to true gain is as shown in Table IV., being as much negative as it was positive in Table II. Now apply the chance error, and the relation between initial obtained scores and gain as computed from the imperfect measures is as shown in Table V., more negative than the real relation is.

In these illustrations the average gain is zero; but it could have been plus any amount or minus any amount without

altering the slopes of the relation lines, which are what concern us.

TABLE II

A REAL RELATION BETWEEN INITIAL AND EVENTUAL ABILITY AND BETWEEN INITIAL ABILITY AND GAIN

Perfect Measure I

Perfect Measure II	1	2	3	4	5	6	7	8	9	10	11
−3	1										
−2		1									
−1											
0		1	1								
1		1	2	1							
2			2	1	2						
3			1	4	3						
4				4	5	3					
5				2	4	3					
6					3	7	4				
7					1	4	5				
8						3	5	6			
9							3	3	2		
10							1	2	1		
11								1	1	2	
12									1	1	
13									1		
14											1

Perfect Measure I

Real Gain (II–I)	1	2	3	4	5	6	7	8	9	10	11
−4	1	1									
−3			1	1	2						
−2		1	2	1	3	3					
−1		1	2	4	5	3	4				
0			1	4	4	7	5	6	2		
+1				2	3	4	5	3	1	2	
+2					1	3	3	2	1	1	
+3							1	1	1		1
+4									1		

The past estimates of the relation of initial ability to improvement [1] are then all underestimates. Their different results for different functions may be largely due to the different amounts of error in the measurements of these functions. For example, Chapman ('14, p. 39) found the correlation between initial score and improvement in addition to be .39 (.44 and .34 by two methods) but that in opposites to be −.14 (−.09 and −.19 by two methods). The correlation of initial with final score was, however, .96 for addition and only .59 for opposites, and the correlations between initial

[1] Such as those of Thorndike ('08, '10), Whitley ('11), Wells ('12), Kirby ('13), and Chapman ('14).

TABLE III

THE EFFECT OF APPLYING A CHANCE ERROR TO THE MEASURES OF TABLE II

Imperfect Measure I

Imperfect Measure II.	1	2	3	4	5	6	7	8	9	10	11
−3		1									
−2											
−1	1										
0	1	1		2							
1			1					1			
2		1	1	3	1	1	1	1			
3		1		1	1	2	1		1		
4	1	1		3	1	1	2			1	
5			2	1	4	4					1
6			2	2	2	2	1	1		1	
7				3	1	4		3			
8				1		2	1		2	1	
9				1	2		1	2	2		
10						1	4	1	2	1	
11									1	1	
12						1			2		1
13									1		
14											1

Imperfect Measure I

Apparent Gain (II–I)	1	2	3	4	5	6	7	8	9	10	11
−7								1			
−6								1	1	1	1
−5		1					1				
−4				2		1	1			1	
−3					1	2	2				
−2	1	1	1	3	1	1		1		1	
−1	1		1	1	1	4	1	3	2		
0		1		3	4	2			2	1	
1		1		1	2	4	1	2	2	1	1
2		1	2	2	1	2	1	1	1		
3	1		2	3			4		2		1
4				1	2	1			1		
5				1							
6						1					

TABLE IV

A REAL RELATION BETWEEN INITIAL ABILITY AND GAIN

Measure I

Real Gain (II–I)	1	2	3	4	5	6	7	8	9	10	11
−4										1	
−3							2	1	1		
−2						3	3	1	2	1	1
−1					4	3	5	4	2	1	
0			2	6	5	7	4	4	1		
1		2	1	3	5	4	3	2			
2		1	1	2	3	3	1				
3	1		1	1	1						
4			1								

score and next to initial score were .97 and .83 respectively. Correction for the larger error in the case of the opposites would raise the −.14 notably.

TABLE V

THE EFFECT OF APPLYING A CHANCE ERROR TO THE MEASURES WHENCE TABLE IV WAS DERIVED

Apparent Gain (II–I) \ Imperfect Measure I	1	2	3	4	5	6	7	8	9	10	11
−8											1
−7										1	
−6							1				
−5								2		1	
−4								2	2	1	
−3					2				2		1
−2						6	3	1	2		1
−1				1		3	2	2	3	2	
0		1		2	3	3	4	1			
1		1	1	5	2	2		1	2		
2				6	3	3	1				
3	3	2	2	2	1						
4			1	1							
5		1	1		1	1					
6			1								

In the case of the gains in mental age made during a time *t*, say for a group of equal chronological age, the error acts of course to make those of higher than average mental ages have smaller apparent gains than they really had, while those of lower than average mental ages have larger apparent gains than they really had. Thus the relation of gain in mental age to initial mental age in such a group seems less positive than it is in reality.

The initially high I.Q.'s will therefore appear to fall more than they do in reality and the initially low I.Q.'s to rise more than they do in reality. For example, in the data of Rugg and Colloton shown in Table VI. the apparent tendency is for initial I.Q. to correlate with gain in I.Q. by −.43 and for each difference of 1 point in initial I.Q. to produce a difference of −.19 point in gain in I.Q. The careless thinker is tempted to fancy that after five or six years, these children will all become alike in I.Q.! Yet the correlation between the two I.Q.'s for an individual is .84 for this series of pairs.

We are working to discover the equation connecting the relation of obtained initial ability with obtained gain, the

unreliability of the measures, and the true facts, especially for for cases where the self-correlation is .60 or higher. Until the equation is available, the best practice is to correlate not against initial ability, but against the average of initial and final ability, so that the minus and plus distortions cancel one another. This is also a useful way to realize the general effect of the error.

TABLE VI

THE OBTAINED RELATION BETWEEN I. Q. AT FIRST TEST AND GAIN IN I. Q. IN SECOND TEST OVER FIRST

		I. Q. at First Test														
		85	90	95	100	105	110	115	120	125	130	135	140	145	150	155
Gain in	−15									1		1				1
I. Q.	−10								3	1	1					1
Sec.	− 5					2	2	3	2	1	1	1	4		1	1
Test	0		2	2	4	3	7	16	3	7	3	4	3		1	
over	+5		1	2	8	3	5	5	3	1	1		1	1		
First	+10	1		1	3	2	2		1	1	1					
Test	+15			1		2	3	2	1	1						

The application of imperfection in the measures to studies of the relation of fatigue or distraction or interest to initial ability is obvious.

Consider now the effect of imperfect measures upon the comparison of groups in respect of gain when the groups are of different initial ability. Returning to our first illustration, let two groups be really alike in all respects of the illustration of Table I., except that one group has an average initial ability of 6, the other of 9. We shall then have differences as follows when those of equal initial scores are compared.

	Initial Score													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Group I.	1.33	1.20	1.67	1.06	.83	− .17	− .64	−1.67	−.55	−2.80	−2.33			
Group II.				1.33	1.20	1.67	1.06	.83	−.17	− .64	−1.67	−.55	−2.80	−2.33
n for group I.	3	5	6	17	12	18	11	9	11	5	3			
n for group II.				3	5	6	17	12	18	11	9	11	5	3
Superiority of group II.	—	—	—	.27	.37	1.94	1.70	2.50	.38	2.16	.66	—	—	—

The average apparent superiority in gain of members of group II. over members of group I. of equal initial ability is

thus from 1.1 to1.3 according to the way the results are weighted, though in reality the two groups gained identically. When the two groups are alike in variability and in the self-correlation of the measure,[1] the correction for this apparent spurious superiority is very simple. It is to subtract $(1-r)d$ from the apparent superiority of group II. where r is the self-correlation of the measure for the groups in question, and where d = (mean of group II) − (mean of group I.). If the individuals had been paired for initial score as by taking

3	from	each	group	of	initial	score	4,
5	"	"	"	"	"	"	5,
6	"	"	"	"	"	"	6,
11	"	"	"	"	"	"	7, and so on,

the correction required would be the same.

The facts reported here add to the evidence that studies of causal factors in psychology will be greatly simplified by extending the measurements of each individual concerned until a self-correlation of at least .95 is attained.

List of References

Chapman, J. C., '14. Individual Differences in Ability and their Correlations. Teachers College, Columbia University, Contributions to Education, No. 63.

Kelley, T. L., '19. Measurement of Overlapping. *Journal of Educational Psychology*, Vol. X., No. 9.

Kelley, T. L., '23. Statistical Method.

Kirby, T. J., '13. Practice in the Case of School Children. Teachers College, Columbia University, Contributions to Education, No. 58.

Thorndike, E. L., '08. The effect of Practice in the Case of Purely Intellectual Function. *American Journal of Psychology*, Vol. 19, pp. 374–384.

Thorndike, E. L., '10. Practice in the Case of Addition. *American Journal of Psychology*, Vol. 21, pp. 483–486.

Thorndike, E. L., '13. An Introduction to the Theory of Mental and Social Measurements. (Revised edition.)

Whitley, M. T., '11. An Empirical Study of Certain Tests for Individual Differences. *Archives of Psychology*, No. 19.

[1] If the two groups are approximately alike in variability and in the self-correlation of the measure, an approximate correction can be made by using the mean of the two self correlations.

A COLOR PREFERENCE SCALE FOR ONE THOUSAND WHITE CHILDREN

BY THOMAS R. GARTH

University of Denver

In a previous article [1] the author presents a color preference scale for 560 Whites along with one for 559 Full Blood Indians. The number of Whites was thus limited in order to avoid results that might come of disparity in size of the groups studied. But it was felt that a scale for white color preference derived from a larger group, one thousand subjects, would be worth obtaining, and it is offered herewith with some effort at analysis of the same. Judgment scales have an advantage over mere statements of sequence in terms of central tendencies in the study of color preference primarily because the position of the colors in a sequence can be determined and their relative 'pulls' measured. But there is also some advantage in presenting these facts of relative 'pull' graphically before the eye. Things which seem to leave in consciousness a sense of difference when compared, though stubbornly refusing to be definitely measured in other ways, thus submit to rather satisfactory quantitative determination.

The task here is (1) to present a color preference scale for Whites, (2) to interpret somewhat the results, (3) to determine any sex differences, (4) to investigate any indications of native tendency as influenced by nurture.

The Sequence of Central Tendencies

In order to approach the study of the scale proper it was thought well for purposes of elucidation to present the data for the one thousand judges in the form of central tendencies, as averages and medians, and the measure of the per cent of the frequencies of a color attaining and exceeding the median of the color preceding it. See Table I.

[1] Garth, Thomas R., 'The Color Preferences of Five Hundred and Fifty-nine Full-Blood Indians,' Jour. of Exp. Psy., Vol. V., 1922, pp. 392–418.

It will be seen that the ranks of the averages and of the medians are the same practically, excepting where the averages of Red and Green are identical. The variabilities are high, but the overlapping of one distribution on another is worth noticing, for it indicates that the sequences are tending to be real excepting in the case of Red and Green where it is only 52 per cent. and Orange and Violet where it is only 56 per cent. The greatest overlapping is in the case of Green and Blue. We may say that this handling of the data indicates that for these subjects it was somewhat tending to be a matter of indifference whether Green went above Red and a matter of somewhat less indifference whether Orange went above Violet, but that they were fairly well decided in the cases of Violet and Red, Yellow and Orange, and Yellow and White respectively. The case was one of absolute surety in the matter of Blue and Green.

TABLE I

SHOWING THE CENTRAL TENDENCIES OF POSITIONS ASSIGNED BY THE JUDGES, THE SEQUENCES, AND OVERLAPPING

Color	Ave.	A. D.	P. E.	Med.	Rank		Overlapping [1]
					Ave.	Med.	
Blue	2.5	1.3	1.1	2.2	1	1	
Green	3.6	1.6	1.4	3.8	2	2	(Blue) 83%
Red	3.6	1.8	1.5	3.9	2	3	(Green) 52%
Violet	4.1	1.6	1.4	4.6	3	4	(Red) 64%
Orange	4.2	1.5	1.3	4.8	4	5	(Violet) 56%
Yellow	4.6	1.4	1.2	5.4	5	6	(Orange) 63%
White	5.5	1.7	1.4	6.9	6	7	(Yellow) 69%

[1] Per cent. of one color attaining and exceeding the median of the preceding color.

THE COLOR PREFERENCE SCALE OF ONE THOUSAND WHITE CHILDREN

Attention is called in the first place to Table II. which shows the scale values for the positions assigned to seven colors by one thousand White school children. The experiment, elsewhere described (*ibid.*), was made by having the subjects arrange in order of preference seven color discs of White, Yellow, Orange, Violet, Green, Red, and Blue, all made of the Milton Bradley Company's 'standard' colored

papers. The discs were one half inch in diameter and were mounted on white cards 1½ by ¾ inches. Records were made of the ranks given the colors by each subject and then data were handled so as to give the scale herewith presented. Figure 1 shows a graphic representation of the same color preference scale for the one thousand judges.

The sequence of the colors, it will be seen, is, running from an arbitrary zero value to a highest value of two hundred four, White, Yellow, Orange, Violet, Red, Green, Blue. The reader will note the exessive fondness for Blue and the light esteem in which Yellow is held by these white subjects. Also, observe that Yellow is about as far above the zero White as its complementary color Blue is above Green, that the complementary colors Red and Green have nearly the same values as have the likewise complementary colors Orange and Violet. If we take the value of Yellow as an index of its 'pull,' we may say in numerical terms that Orange and Violet 'pull' one and one half times as strong, and Red and Green almost twice (1.8) as strong as Yellow, whereas Blue 'pulls' almost three times as strong as that color. Blue "pulls" twice as strong, approximately, as Orange or Violet, and about one and one half (slightly more) times as much as Red and Green.

The writer has no theory to offer for the fact that Blue is the most preferred color and Yellow is next to the zero White in preference value. The[1] color preference scale does not correlate with the physical facts of color as for instance as respects the different wave lengths. The all but shortest wave-length is here most pleasing, as Blue, and the next shortest, as Green, next in pleasing value, while the next color, Red, standing close beside the preceding in pleasing value, has the longest wave-length in the series. Red is third, almost second, in pleasing value, while Yellow with the next wave-length to Green is far below Red and Orange with longer wave-lengths, while Violet with shortest wave-length of all comes between Orange and Red.

Fèrè[2] ('Sensation et Mouvement') found in a study with

[1] Woodworth, R. S., 'The Psychology of Light,' Reprinted from the *Transaction of the Illuminating Engineering Society*, May, 1916, p. 467.

[2] *Ibid.*

the dynamometer, that red light served to stimulate his subjects more than any other color of light, so that it seemed to facilitate the contraction of the muscles. Next came orange light, green, yellow, and blue lights in a series of degrees of ability to stimulate the subject. But the color preference scale does not correlate with the stimulation series.

If [1] again we should try to interpret the results of the preference scale for Whites in terms of color symbolism, the effort is likewise abortive, starting out well but failing in the end, as Blue we placed first because it symbolizes peacefulness, depth, solidity, coolness; Green is next, being cool also, slightly less calm, and suggesting less of depth, hopefulness; but it is followed by Red the most exciting of colors, suggesting warmth and passion, strength; and at the end of the scale is the joyous Yellow which is said to be warm and uplifting.

Table II

Showing a Color Preference Scale for One Thousand White Children

Color	Scale Value
White	0
Yellow	.69
Orange	1.00
Violet	1.07
Red	1.22
Green	1.26
Blue	2.04

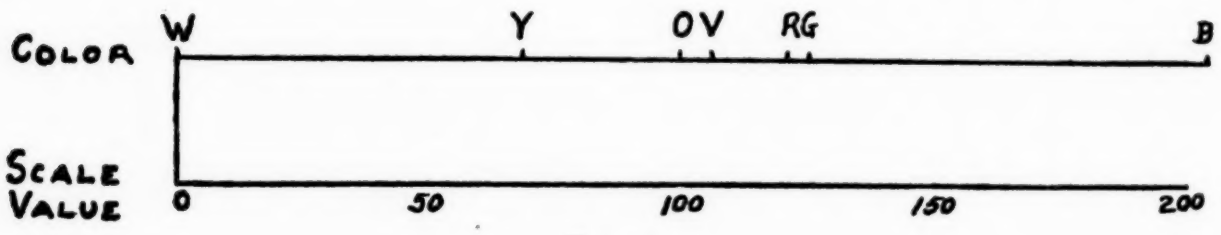


Fig. 1

Sex Differences in Color Preference

If we examine Table III., we find the color preference scales for Males and Females. The two sexes agree in placing White last, Yellow next to last, and Blue first. The males are fonder of Red than are the females, the former placing it second best, while the latter place it fifth. In both scales Blue has about the same value, 2.29 for males and 2.24 for females. Since the sex groups are composed of about equal numbers of

[1] Gordon, Kate, 'Esthetics,' pp. 146 ff.

individuals, it seems legitimate to make the comparisons of the values assigned to the several colors given in Table IV. Here, we find the ratio of the value assigned by females to that assigned by males. It is seen that there is most agreement in respect to Blue, first, then Green, then Orange. There is least agreement in the case of Red, next Orange, and there is absolute agreement with respect to Yellow. These sex differences as found here do not agree with some of the previous findings of studies in color preference, as for instance the study made by Jastrow in 1893. By eliminating Blue temporarily, we may say that the males have clearer discrimination of color preference than the females. Green has, in the case of the males, 2.4 times the 'pull' of Yellow and for the females only 1.9 times that of Yellow. In the case of Violet in terms of its relative 'pull' to that of Yellow, for males it is 2.3 and for females 1.2. We give these ratios in terms of Yellow below:

For Males

Yellow	Orange	Violet	Green	Red	Blue
1	2.0	2.3	2.4	2.9	3.0

For Females

Yellow	Red	Violet	Orange	Green	Blue
1	1.1	1.2	1.4	1.9	3.0

TABLE III

SHOWING COLOR PREFERENCE SCALES FOR WHITE MALES AND FEMALES

Scale for 498 Males		*Scale for 502 Females*	
Color	Scale Value	Color	Scale Value
White	.0	White	.0
Yellow	.74	Yellow	.74
Orange	1.43	Red	.78
Violet	1.73	Violet	.89
Green	1.84	Orange	1.04
Red	2.18	Green	1.42
Blue	2.29	Blue	2.24

THE INFLUENCE OF EDUCATION ON COLOR PREFERENCE

We call attention here to Table V. which gives the different scales for the various grades in school beginning with the first, and running up to the most educated group which is composed of students from the eighth, ninth, and tenth grades. This makes in all eight educational groups.

In the lowest educational group, the first grade group, we find Red is slightly better than Blue, but only slightly so, and it is interesting to note that this is a very temporary advantage for in the second grade scale it falls as far below Blue as it was previously above it. From this point on, it has a tendency to be less and less liked by the successive educational groups, though the fall is not regular, until it reaches an extremely low point on the scale of the last educational group. The data for the other colors excepting Blue have the same appearance so that when one looks at Table VI, which gives the color values for two successive grades averaged one with the other, and the curves of Figure 2, one sees quite plainly that there is a tendency for the colors generally, with the exception of the color Blue, to experience a drop in color appreciation, allowing for a slight rise in the case of Green. But the color Blue is persistently high in the scale regardless of the educational group. The Red, Orange, and Yellow lines fall rather precipitately, but of the three the fall is greatest for the Red line. The Green and Violet lines experience some fluctuations. However, the line for Blue is almost horizontal.

In the light of the conditions of the experiment the conclusion seems justifiable that the subjects experience a suppression of response to all colors excepting Blue with the increase of education.

TABLE IV

SHOWING RELATIVE PULL OF THE COLORS ON MALES AND FEMALES

Color value for females divided by the value for males.

Yellow	100%
Orange	73 "
Violet	51 "
Green	75 "
Red	36 "
Blue	97 "

SHOWING COLOR PREFERENCE SCALES FOR THE SEVERAL GRADES

TABLE V

First Grade 158 Judges		*Second Grade 131 Judges*	
Color	Scale Value	Color	Scale Value
White	.0	White	.0
Yellow	1.20	Yellow	1.30
Violet	1.42	Orange	1.75
Green	1.49	Green	1.86
Orange	1.60	Violet	1.90
Blue	1.94	Red	2.24
Red	1.98	Blue	2.28

Third Grade—136 Judges

Color	Scale Value
White	.0
Violet	.82
Yellow	.86
Orange	1.47
Red	1.85
Green	1.89
Blue	2.58

Fourth Grade [1]*—115 Judges*

Color	Scale Value
White	.0
Violet	.65
Yellow	.74
Orange	1.22
Red	1.52
Green	1.71
Blue	1.90

Fifth Grade—115 Judges

Color	Scale Value
White	.0
Yellow	.53
Orange	.72
Violet	.94
Red	1.39
Green	1.44
Blue	1.71

Sixth Grade—117 Judges

Color	Scale Value
White	.0
Yellow	.49
Orange	.98
Violet	1.09
Red	1.09
Green	1.39
Blue	2.39

Seventh Grade—120 Judges

Color	Scale Value
White	.0
Yellow	.0
Red	.41
Violet	.56
Orange	.60
Green	1.01
Blue	2.06

Eight, Ninth and Tenth Grades—108 Judges

Color	Scale Value
White	.0
Yellow	.22
Red	.48
Orange	.48
Violet	.78
Green	1.00
Blue	1.94

TABLE VI

SHOWING AVERAGE OF VALUES ASSIGNED FOR RESPECTIVE EDUCATIONAL GROUPS

Grade	Yellow	Violet	Green	Orange	Red	Blue
7-10	22	67	105	54	45	200
5-6	51	102	142	85	124	205
3-4	80	72	180	134	169	224
1-2	125	166	168	168	211	211

SUMMARY

1. The color preference sequence of the seven colors used in the experiment is for these one thousand white subjects running from most preferred to least preferred: Blue, Green, Red, Violet, Orange, Yellow, White.

2. The measures of overlapping indicate that the differences are significant excepting in the case of Red and Green, and Violet and Orange, two pairs of complementary colors.

[1] The scales for grades four through six have been printed before in the articl referred to above (*Ibid.*).

3. In terms of 'pull' as derived from the scale calling the 'pulling power' of Yellow unity, we have the following:

Yellow	Orange and Violet	Red and Green	Blue
1	1.5	1.8	3.0

4. Blue 'pulls' twice as hard as Orange and Violet, three times as hard as Yellow, and one and two thirds times as hard as Red or Green. Relatively speaking it has an exceedingly strong 'pull.'

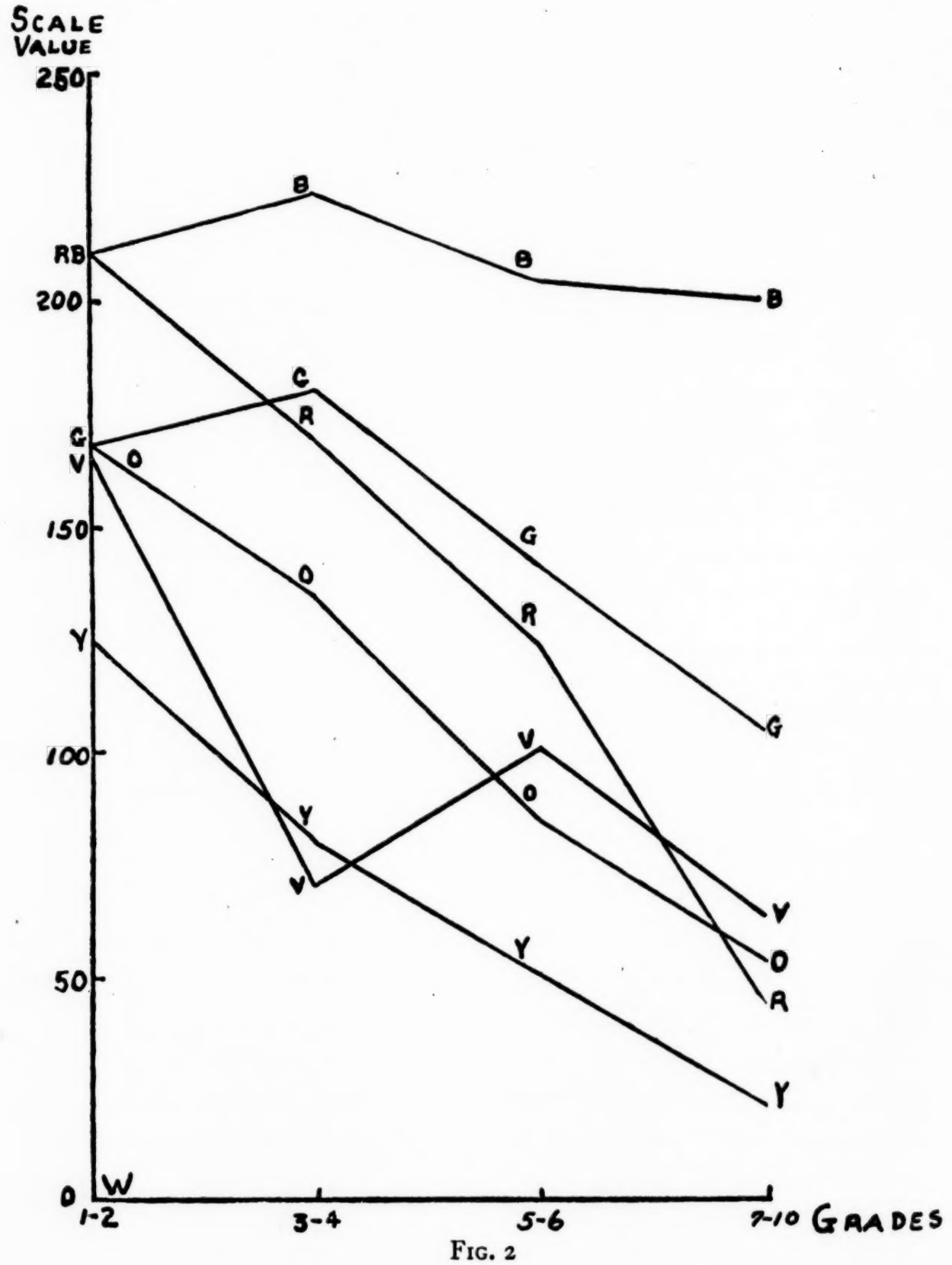


FIG. 2

5. All colors 'pull' harder on males than they do on females, excepting in the case of Yellow and Blue where the ratios of scale value on male values as a base are respectively 100 and 97. The greatest disagreement between the sexes is in the case of Red which for males has a high scale value and for the females a low scale value.

6. It is impossible with this material to separate the effects of education and age, but when we arrange our scales in an educational sequence, we get a tendency for Blue to be ever held in highest esteem excepting in the case of the First Grade subjects where Red is slightly preferred to it, but only slightly. Red falls precipitately with education, losing three fourths of its 'pull' finally. Orange loses nearly two thirds of its original 'pull' when we compare the least educated group scale values (First and Second) with the most educated group scale values (Seventh to Tenth Grades). While Green and Violet have a tendency to fluctuate they too finally fall with increased education, but Violet more than Green.

7. Education seems to produce a tendency toward a suppression of preference for all colors excepting Blue in the case of White subjects, as measured in terms of scale values.

PUBLISHER'S NOTICE

The cost of printing has increased considerably since our present subscription rates came into effect in 1920. We have delayed making additional changes, in the hope that the peak was reached. On the contrary there has been a further rise in printing costs within the past few months. We are therefore compelled, most reluctantly, to fix a new scale of rates for all our publications. The revised list appears on the second cover page. Subscribers will please note the advantageous combination rates offered.

The new prices will take effect July 1.

PSYCHOLOGICAL REVIEW COMPANY